

Db 181 GGGTCAGTTATGTGCAAGAACCTTTGGCGCTTCTTGCGCAGCAAGTACCATGGGGTT 240
Qy 241 GTGCTCGAAGTGTCTACCGAGAGACATCATGCAAGCGAAAGATGAGGGCTTTAGCTGAGCA 300
Db 241 GTGCTCGAAGTGTCTACCGAGAGACATCATGCAAGCGAAAGATGAGGGCTTTAGCTGAGCA 300
Qy 301 AGCCACTCAGGCTGCTCAGGCGACATCTGSCACAGCTGCTGTTAGAGCCCCCGCTCC 360
Db 301 AGCCACTCAGGCTGCTCAGGCGACATCTGSCACAGCTGCTGTTAGAGCCCCCGCTCC 360
Qy 361 TGTACATGAGACCAAGCTCACAATGCGAGTTGAGAGAACATGATTGTGCCCATCAATC 420
Db 361 TGTACATGAGACCAAGCTCACAATGCGAGTTGAGAGAACATGATTGTGCCCATCAATC 420
Qy 421 TTCCAGCTATCAACAAGACTGTTACCCCGCTGAGTGCCTCTCAGGCGATGAATC 480
Db 421 TTCCAGCTATCAACAAGACTGTTACCCCGCTGAGTGCCTCTCAGGCGATGAATC 480
Qy 481 CTCTATCGCAGCTCCCTCTAGACCCGAGCCCAATGCAATGCGATCTTGCGAGGAAGTGT 540
Db 481 CTCTATCGCAGCTCCCTCTAGACCCGAGCCCAATGCAATGCGATCTTGCGAGGAAGTGT 540
Qy 541 TGGATTGACAGGATTTAAGTGTGCTGTGGCAACCTTAATGCGCTTTACATCGGTAATC 600
Db 541 TGGATTGACAGGATTTAAGTGTGCTGTGGCAACCTTAATGCGCTTTACATCGGTAATC 600
Qy 601 GGACAAACACACTTGTGACATATGACTACAAAGCCGAGGGCAGAAAGCATTTGCAAAAGC 660
Db 601 GGACAAACACACTTGTGACATATGACTACAAAGCCGAGGGCAGAAAGCATTTGCAAAAGC 660
Qy 661 TAACTCTTGTGCTGCGGCGAGAGGTTGTCAAGTTTGTATGAGCATCCGTTAACTTTT 720
Db 661 TAACTCTTGTGCTGCGGCGAGAGGTTGTCAAGTTTGTATGAGCATCCGTTAACTTTT 720
Qy 721 CTGCCGACGATTAGGCTTCATACATTGAGTAATCTTACATCTTTCTTTATGAGAG 780
Db 721 CTGCCGACGATTAGGCTTCATACATTGAGTAATCTTACATCTTTCTTTATGAGAG 780
Qy 781 AGCGAGTCGCATCAAGAGCTCGCC 804
Db 781 AGCGAGTCGCATCAAGAGCTCGCC 804

RESULT 2
AX288141 804 bp DNA linear PAT 21-NOV-2001
DEFINITION Sequence 12 from Patent WO0177311.
ACCESSION AX288141
VERSION AX288141.1 GI:17049843
KEYWORDS
SOURCE .
ORGANISM Physcomitrella patens
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella.
REFERENCE 1
da Costa Silva,O., Bohner,H.J., Van Thiel,N., and Chen,R.
Transcription factor stress-related proteins and methods of use in
Plants
Patent: WO 0177311-A 12 18-OCT-2001;
JOURNAL BASF Plant Science GmbH (DE)
FEATURES
source 1..804
/organism="Physcomitrella patens"
/mol_type="unassigned DNA"
/db_xref="taxon:3218"

ORIGIN
Query Match 100.0%; Score 804; DB 6; Length 804;
Best Local Similarity 100.0%; Pred. No. 5,6e-236;
Matches 804; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 ATCCGGGACACAGTCCCGCTTATGTGTGTCTATTAGTGTGTTGTTGCAAGCTGAAGC 60
1 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 1 ATCCGGGACACAGTCCCGCTTATGTGTGTCTATTAGTGTGTTGCTC 60
Qy 61 CTGAGCGAGATTGTCAGGATTTTCTCATACGCTTCTGATTAGAAAGA 120
Db 61 CTGAGCGAGATTGTCAGGATTTTCTCATACGCTTCTGATTAGAAAGA 120
Qy 121 TAGTGTGTTAAAGATGCGCACCGAGCGTGTCTCAGAGAGACCTCG 180
Db 121 TAGTGTGTTAAAGATGCGCACCGAGCGTGTCTCAGAGAGACCTCG 180
Qy 181 GGGTCAGTTATGTGCAAGAACTTTGGCGCTTCTTGCGCAGCCAACT 240
Db 181 GGGTCAGTTATGTGCAAGAACTTTGGCGCTTCTTGCGCAGCCAACT 240
Qy 241 GTGCTCGAAGTGTCTACCGAGAGACATGATGCAAGCGAAATGAGCGCT 300
Db 241 GTGCTCGAAGTGTCTACCGAGAGACATGATGCAAGCGAAATGAGCGCT 300
Qy 301 AGCCACTCAGGCTGCTCAGGCGACATCTGSCACAGCTGCTGTTAG 360
Db 301 AGCCACTCAGGCTGCTCAGGCGACATCTGSCACAGCTGCTGTTAG 360
Qy 361 TGTACATGAGACCAAGCTCACAATGCGAGTTGAGAGAACATGATTGTG 420
Db 361 TGTACATGAGACCAAGCTCACAATGCGAGTTGAGAGAACATGATTGTG 420
Qy 421 TTCCAGCTATCAACAAGACTGTTACCCCGCTGAGTGCCTCTCAG 480
Db 421 TTCCAGCTATCAACAAGACTGTTACCCCGCTGAGTGCCTCTCAG 480
Qy 481 CTCTATCGCAGCTCCCTCTAGACCCGAGCCCAATGCAATGCGATCTTGC 540
Db 481 CTCTATCGCAGCTCCCTCTAGACCCGAGCCCAATGCAATGCGATCTTGC 540
Qy 541 TGGATTGACAGGATTTAAGTGTGCTGTGGCAACCTTAATGCGCTTTA 600
Db 541 TGGATTGACAGGATTTAAGTGTGCTGTGGCAACCTTAATGCGCTTTA 600
Qy 601 GGACAAACACACTTGTGACATATGACTACAAAGCCGAGGGCAGAAAGC 660
Db 601 GGACAAACACACTTGTGACATATGACTACAAAGCCGAGGGCAGAAAGC 660
Qy 661 TAACTCTTGTGCTGCGGCGAGAGGTTGTCAAGTTTGTATGAGCATCC 720
Db 661 TAACTCTTGTGCTGCGGCGAGAGGTTGTCAAGTTTGTATGAGCATCC 720
Qy 721 CTGCCGACGATTAGGCTTCATACATTGAGTAATCTTACATCTTTCTTC 780
Db 721 CTGCCGACGATTAGGCTTCATACATTGAGTAATCTTACATCTTTCTTC 780
Qy 781 AGCGAGTCGCATCAAGAGCTCGCC 804
Db 781 AGCGAGTCGCATCAAGAGCTCGCC 804

RESULT 3
AR452942 1531 bp DNA linear 2004
DEFINITION Sequence 4 from patent US 6677504.
ACCESSION AR452942
VERSION AR452942.1 GI:42685089
KEYWORDS
SOURCE .
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1531)
da Costa e Silva,O., Bohner,H.J., Van Thiel,N., a
Transcription factor stress-related proteins and me
Patent: US 6677504-A 4 13-JAN-2004;
JOURNAL BASF Plant Science GmbH; Ludwigshafen;
DEX; Location/Qualifiers

```
source 1..1531
/organism="unknown"
/mol_type="genomic DNA"

Query Match 96.6%; Score 776.4; DB 6; Length 1531;
Best Local Similarity 99.7%; Pred. No. 1.8e-227;
Matches 788; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

ORIGIN
Query Match 96.6%; Score 776.4; DB 6; Length 1531;
Best Local Similarity 99.7%; Pred. No. 1.8e-227;
Matches 788; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

8 GCACCAAGTCCCGCTTAGTGTGTGTCATTAGTGTGTCAGAGCTTGAC 67
529 GCACCAAGTCCCGCTTAGTGTGTGTCATTAGTGTGTCAGAGCTTGAC 588
68 GAGATTGACAGATTTTCTCATACGCTTCTGATTAAGAAAGATACCCCTTATTAGCTG 127
589 GAGATTGACAGATTTTCTCATACGCTTCTGATTAAGAAAGATACCCCTTATTAGCTG 648
128 TTAAGATGGCCAGCAGAGCTGTGTCAGAGACGACCTCGAGGCGCTGAGGGTCCA 187
649 TTAAGATGGCCAGCAGAGCTGTGTCAGAGACGACCTCGAGGCGCTGAGGGTCCA 708
188 GTTATGTGCAAGAACCTTTGCGGCTTCTTGGCAGCCAACTACATGGGGTTGTGCTG 247
709 GTTATGTGCAAGAACCTTTGCGGCTTCTTGGCAGCCAACTACATGGGGTTGTGCTG 768
248 AAGTGTACCGAGAGACATGTCAGAGGAAAGATACGGCTTTAGCTGACAGCACT 307
769 AAGTGTACCGAGAGACATGTCAGAGGAAAGATACGGCTTTAGCTGACAGCACT 827
308 CAGGCTGTCAAGGCGACATCTGCAAGCTGTGCTTCAAGCCCGCTCTGTACAT 367
828 CAGGCTGTCAAGGCGACATCTGCAAGCTGTGCTTCAAGCCCGCTCTGTACAT 887
368 GAGACCAAGCTCAATGCGAGGTTGAGAGAACATATGTCGCCATCAATCTTCCAGC 427
888 GAGACCAAGCTCAATGCGAGGTTGAGAGAACATATGTCGCCATCAATCTTCCAGC 947
428 TATCAACAAGACCTGTGTTACCCCGCTGACGCTGCCCCCTCAGGCAATGCTTATC 487
948 TATCAACAAGACCTGTGTTACCCCGCTGACGCTGCCCCCTCAGGCAATGCTTATC 1007
488 GAGCTCCCTCTAGACCCGAGCCCAATGCAATGCGGATTTGAGAGAGGCTGTGGAATTG 547
1008 GAGCTCCCTCTAGACCCGAGCCCAATGCAATGCGGATTTGAGAGAGGCTGTGGAATTG 1067
548 ACAGGATTTAAGTGTGCTGTGCAACTCTTACTGCGCTTTAATCGATCTCGACAA 607
1068 ACAGGATTTAAGTGTGCTGTGCAACTCTTACTGCGCTTTAATCGATCTCGACAA 1127
608 CACACTTGCACTATGATCAAAAGCCGAGGCGAGAAAGCAATGCGAAAGCTAATCT 667
1128 CACACTTGCACTATGATCAAAAGCCGAGGCGAGAAAGCAATGCGAAAGCTAATCT 1187
668 CTGTGCGTGGCGAGAGGTTGCAAGTTTATGAGCAATCGTTAAGCTTTTTCGCGA 727
1188 CTGTGCGTGGCGAGAGGTTGCAAGTTTATGAGCAATCGTTAAGCTTTTTCGCGA 1247
728 CGATTAGGCTTCATACATTTGATGTAATCTTCTTCTTTATCGAGAGAGCAAGT 787
1248 CGATTAGGCTTCATACATTTGATGTAATCTTCTTCTTTATCGAGAGAGCAAGT 1307
788 CGCATCAAGA 797
1308 CGCATCAAGA 1317

RESULT 4
AX288133 1531 bp DNA linear PAT 21-NOV-2001
LOCUS AX288133
DEFINITION Sequence 4 from Patent WO0177311.
ACCESSION AX288133
VERSION AX288133.1 GI:17049835
KEYWORDS
```

```
SOURCE Physcomitrella patens
ORGANISM Physcomitrella patens
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta
AUTHORS Bryopsida; Funariidae; Funariales; Funariaceae; Phy
1 da Costa Silva O., Bohner H.J., Van Thiel N., and
TITLE Transcription factor stress-related proteins and me
JOURNAL Patent: WO 0177311-A 4 18-OCT-2001;
FEATURES
source BASF Plant Science GmbH (DE)
location/Qualifiers
1..1531
/organism="Physcomitrella patens"
/mol_type="unassigned DNA"
/db_xref="taxon:3218"

ORIGIN
Query Match 96.6%; Score 776.4; DB 6; Length 1
Best Local Similarity 99.7%; Pred. No. 1.8e-227;
Matches 788; Conservative 0; Mismatches 1; Indels 1;

8 GCACCAAGTCCCGCTTAGTGTGTGTCATTAGTGTGTCAGAGCTTGAC 67
529 GCACCAAGTCCCGCTTAGTGTGTGTCATTAGTGTGTCAGAGCTTGAC 588
68 GAGATTGACAGATTTTCTCATACGCTTCTGATTAAGAAAGATACCC 127
589 GAGATTGACAGATTTTCTCATACGCTTCTGATTAAGAAAGATACCC 648
128 TTAAGATGGCCAGCAGAGCTGTGTCAGAGACGACCTCGAGGCGCT 187
649 TTAAGATGGCCAGCAGAGCTGTGTCAGAGACGACCTCGAGGCGCT 708
188 GTTATGTGCAAGAACCTTTGCGGCTTCTTGGCAGCCAACTACATG 247
709 GTTATGTGCAAGAACCTTTGCGGCTTCTTGGCAGCCAACTACATG 768
248 AAGTGTACCGAGAGACATGTCAGAGGAAAGATACGGCTTTAGCT 307
769 AAGTGTACCGAGAGACATGTCAGAGGAAAGATACGGCTTTAGCT 827
308 CAGGCTGTCAAGGCGACATCTGCAAGCTGTGCTTCAAGCCCGCT 367
828 CAGGCTGTCAAGGCGACATCTGCAAGCTGTGCTTCAAGCCCGCT 887
368 GAGACCAAGCTCAATGCGAGGTTGAGAGAACATATGTCGCCAT 427
888 GAGACCAAGCTCAATGCGAGGTTGAGAGAACATATGTCGCCAT 947
428 TATCAACAAGACCTGTGTTACCCCGCTGACGCTGCCCCCTCAG 487
948 TATCAACAAGACCTGTGTTACCCCGCTGACGCTGCCCCCTCAG 1007
488 GAGCTCCCTCTAGACCCGAGCCCAATGCAATGCGGATTTGAGAG 547
1008 GAGCTCCCTCTAGACCCGAGCCCAATGCAATGCGGATTTGAGAG 1067
548 ACAGGATTTAAGTGTGCTGTGCAACTCTTACTGCGCTTTAATCG 607
1068 ACAGGATTTAAGTGTGCTGTGCAACTCTTACTGCGCTTTAATCG 1127
608 CACACTTGCACTATGATCAAAAGCCGAGGCGAGAAAGCAATGCG 667
1128 CACACTTGCACTATGATCAAAAGCCGAGGCGAGAAAGCAATGCG 1187
668 CTGTGCGTGGCGAGAGGTTGCAAGTTTATGAGCAATCGTTAAG 727
1188 CTGTGCGTGGCGAGAGGTTGCAAGTTTATGAGCAATCGTTAAG 1247
728 CGATTAGGCTTCATACATTTGATGTAATCTTCTTCTTTATCG 787
1248 CGATTAGGCTTCATACATTTGATGTAATCTTCTTCTTTATCG 1307
```

Oy 788 CGCATCAGA 797
| | | | |
Db 1308 CGCATCAGA 1317

RESULT 5
AC161406

LOCUS
DEFINITION AC161406 112882 bp DNA linear HTG 13-MAY-2005
Medicago truncatula chromosome 2 clone mth2-83p17, *** SEQUENCING
IN PROGRESS ***, 32 unordered pieces.

AC161406
AC161406.1 GI:63987046
HTG: HTGS PHASE1, HTGS ACTIVEFIN.
KEYWORDS Medicago truncatula (barrel medic)
SOURCE Medicago truncatula
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
1 (bases 1 to 112882)
Town,C.D., Tallon,L.T., Arbogast,T., Althoff,R., Rhine,E.,
Moneghun,E., Smith,S.A., Utecherack,T., Feldlyum,T., Koo,H. and
Cheung,F.
Medicago truncatula BAC genomic sequence
Unpublished
2 (bases 1 to 112882)
Town,C.D.
Direct Submission
Submitted (13-MAY-2005) The Institute for Genomic Research, 9712
Medical Center Dr., Rockville, MD 20850, USA
* NOTE: This is a 'working draft' sequence. It currently
* consists of 32 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 10642: contig of 10642 bp in length
10643 10742: gap of unknown length
10743 21453: contig of 10711 bp in length
21454 21553: gap of unknown length
21554 23120: contig of 1567 bp in length
23121 23220: gap of unknown length
23221 28798: contig of 5578 bp in length
28799 28898: gap of unknown length
28899 30125: contig of 1227 bp in length
30126 30225: gap of unknown length
30226 39880: contig of 9655 bp in length
39881 39980: gap of unknown length
39981 41417: contig of 1437 bp in length
41418 41517: gap of unknown length
41518 42809: contig of 1292 bp in length
42810 42909: gap of unknown length
42910 44087: contig of 1178 bp in length
44088 44187: gap of unknown length
44188 47787: contig of 3600 bp in length
47788 47887: gap of unknown length
47888 48891: contig of 1004 bp in length
48891 48991: gap of unknown length
48992 50482: contig of 1491 bp in length
50483 50582: gap of unknown length
50583 51672: contig of 1050 bp in length
51673 51772: gap of unknown length
51773 52840: contig of 1068 bp in length
52841 52940: gap of unknown length
52941 70023: contig of 17083 bp in length
70024 70123: gap of unknown length
70124 71565: contig of 1442 bp in length
71566 71665: gap of unknown length
71666 72803: contig of 118 bp in length
72804 72903: gap of unknown length
72904 74559: contig of 1656 bp in length

* 74560 74559: gap of unknown length
* 74560 79544: contig of 4885 bp in length
* 79545 79644: gap of unknown length
* 79645 90908: contig of 11264 bp in length
* 90909 91008: gap of unknown length
* 91009 92604: contig of 1596 bp in length
* 92605 92704: gap of unknown length
* 92705 94232: contig of 1528 bp in length
* 94233 94332: gap of unknown length
* 94333 95349: contig of 1017 bp in length
* 95350 95449: gap of unknown length
* 95450 100396: contig of 4947 bp in length
* 100397 100496: gap of unknown length
* 100497 101862: contig of 1366 bp in length
* 101863 101962: gap of unknown length
* 101963 103478: contig of 1516 bp in length
* 103479 103578: gap of unknown length
* 103579 105211: contig of 1633 bp in length
* 105212 105311: gap of unknown length
* 105312 106455: contig of 1144 bp in length
* 106456 106555: gap of unknown length
* 106556 107581: contig of 1026 bp in length
* 107582 107681: gap of unknown length
* 107682 110219: contig of 2538 bp in length
* 110220 110319: gap of unknown length
* 110320 111674: contig of 1355 bp in length
* 111675 111774: gap of unknown length
* 111775 112882: contig of 1108 bp in length.

FEATURES
source
1..112882
Location/Qualifiers
/organism="Medicago truncatula"
/mol_type="genomic DNA"
/db_xref="taxon:3880"
/chromosome="2"
/clone="mth2-83p17"
10643..10742
/estimated_length=unknown
21454..21553
/estimated_length=unknown
23121..23220
/estimated_length=unknown
28799..28898
/estimated_length=unknown
30126..30225
/estimated_length=unknown
39881..39980
/estimated_length=unknown
41418..41517
/estimated_length=unknown
42810..42909
/estimated_length=unknown
44088..44187
/estimated_length=unknown
47788..47887
/estimated_length=unknown
48891..48991
/estimated_length=unknown
50483..50582
/estimated_length=unknown
51673..51772
/estimated_length=unknown
52841..52940
/estimated_length=unknown
70024..70123
/estimated_length=unknown
71566..71665
/estimated_length=unknown
72804..72903
/estimated_length=unknown
74560..74559
/estimated_length=unknown
79545..79644
/estimated_length=unknown

gap 90909..91008 /estimated_length=unknown
gap 92605..92704 /estimated_length=unknown
gap 94233..94332 /estimated_length=unknown
gap 95350..95449 /estimated_length=unknown
gap 100397..100496 /estimated_length=unknown
gap 101863..101962 /estimated_length=unknown
gap 103479..103578 /estimated_length=unknown
gap 105212..105311 /estimated_length=unknown
gap 106456..106555 /estimated_length=unknown
gap 107582..107681 /estimated_length=unknown
gap 110220..110319 /estimated_length=unknown
gap 111675..111774 /estimated_length=unknown
gap /estimated_length=unknown
ORIGIN
Query Match 13.1%; Score 105; DB 14; Length 112882;
Best Local Similarity 67.7%; Pred. No. 2e-20;
Matches 147; Conservative 0; Mismatches 70; Indels 0; Gaps 0;
QY 509 CCCAATCGATGCGGATCTTGACAGAGCGCTTGATGACAGATTAAAGTCCTCT 568
DB 84612 CCTAAGCGGTGACGACATGCAATAGCGGGTTGTTGACAGATTATGTCATGT 84671
QY 569 GCGAAGCTTACTGCGCTTACCTCGGTACTCGGACAAACACACTTGCATATGACTAC 628
DB 84672 GGTATCTTACTGCTGTCTGATCCCTACTCGAGAACAGACGACGCCATTGATTAC 84731
QY 629 AAGCCGCGAGGAGAGACGATGCGAAAGCTAATCTCTGTGCGGCCGAGAAAGTT 688
DB 84732 CCACACTGCGAGCGATGATCAATAGCAAAAGCAACCTGTGTGAAAGGCTGAAAGCTT 84791
QY 689 GTCAGTTTGTATGAGCATCCGTTAAGCTTTCTGCC 725
DB 84792 GACAAATCTAGATTGTTGTCTTAAAGTTTCATGTC 84828
RESULT 6
AP007400/c 102653 bp DNA linear HTG 28-DEC-2004
LOCUS AP007400
DEFINITION Lotus corniculatus var. japonicus chromosome 3 clone LjT31F23, ***
SEQUENCING IN PROGRESS ***, 37 unordered pieces.
ACCESSION AP007400
VERSION AP007400.1 GI:56805716
KEYWORDS HTG; HTGS PHASE1.
SOURCE Lotus corniculatus var. japonicus (Lotus japonicus)
ORGANISM Lotus corniculatus var. japonicus
Spermatophyta; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Rosales; eucosydes I; Fabales; Fabaceae; Papilionoideae; Lotaceae;
Lotus.
REFERENCE 1
AUTHORS Kaneko,T., Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
TITLE Structural Analysis of a Lotus japonicus Genome. XI. Sequence
FEATURES AND Mapping of Nine hundred twenty-one TAC Clones
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 102653)
AUTHORS Sato,S.
TITLE Direct Submision
JOURNAL Submitted (26-OCT-2004) Shuei Sato, Kazusa DNA Research Institute,
Department of Plant Gene Research, 2-6-7 Kazusa-kamatarai, Kisarazu,
Chiba, 292-0818, Japan (E-mail:issato@kazusa.or.jp,
URL:htp://www.kazusa.or.jp/, Tel:81-438-52-3935(ex.2337)),

COMMENT

Fax:81-438-52-3934)
* NOTE: This is a 'working draft' sequence. It con
* consists of 37 contigs. The true order of the p
* is not known and their order in this sequence re
* arbitrary. Gaps between the contigs are represent
* runs of N, but the exact sizes of the gaps are u
* This record will be updated with the finished seq
* as soon as it is available and the accession num
* be preserved.
1
772 771: contig of 771 bp in length
872 871: gap of unknown length
1868 1867: contig of 996 bp in length
1968 1967: gap of unknown length
2909 2908: contig of 941 bp in length
3009 3008: gap of unknown length
3973 3972: contig of 964 bp in length
4073 4072: gap of unknown length
5829 5828: contig of 1756 bp in length
5929 5928: gap of unknown length
7171 7170: contig of 1242 bp in length
7271 7270: gap of unknown length
8999 8998: contig of 1728 bp in length
9099 9098: gap of unknown length
10096 10095: contig of 997 bp in length
10196 10195: gap of unknown length
11820 11819: contig of 1624 bp in length
11920 11919: gap of unknown length
13526 13525: contig of 1606 bp in length
13626 13625: gap of unknown length
15096 15095: contig of 1470 bp in length
15196 15195: gap of unknown length
17491 17491: contig of 2296 bp in length
17592 17591: gap of unknown length
19659 19658: contig of 2067 bp in length
19758 19758: gap of unknown length
21397 21397: contig of 1639 bp in length
21497 21497: gap of unknown length
21498 21498: contig of 2378 bp in length
23876 23875: gap of unknown length
23976 23975: gap of unknown length
25652 25652: contig of 1677 bp in length
25752 25752: gap of unknown length
26597 26597: contig of 845 bp in length
26598 26597: gap of unknown length
28740 28740: contig of 2043 bp in length
28840 28840: gap of unknown length
28741 28841: contig of 2777 bp in length
31617 31617: gap of unknown length
31718 31718: contig of 1883 bp in length
33601 33600: contig of 1967 bp in length
33701 33700: gap of unknown length
36667 36667: contig of 2967 bp in length
36767 36767: gap of unknown length
39712 39712: contig of 2945 bp in length
39812 39812: gap of unknown length
39813 39813: contig of 2677 bp in length
42489 42489: gap of unknown length
42589 42589: gap of unknown length
45916 45916: contig of 3327 bp in length
45917 45917: gap of unknown length
46016 46016: gap of unknown length
49063 49063: contig of 3047 bp in length
49064 49064: gap of unknown length
49164 49164: contig of 3081 bp in length
52344 52344: contig of 3081 bp in length
52345 52345: gap of unknown length
53907 53907: contig of 3563 bp in length
55908 55908: gap of unknown length
59274 59274: contig of 3267 bp in length
59375 59375: gap of unknown length
61819 61819: contig of 2445 bp in length
61820 61819: gap of unknown length
61920 61920: contig of 5062 bp in length
66982 67081: gap of unknown length
67082 72120: contig of 5039 bp in length
72121 72220: gap of unknown length
72221 75654: contig of 3434 bp in length
75655 75754: gap of unknown length

* 75755	81137: contig of 5383 bp in length
* 81138	81237: gap of unknown length
* 81238	85715: contig of 4478 bp in length
* 85716	85815: gap of unknown length
* 85816	91795: contig of 5980 bp in length
* 91796	91895: gap of unknown length
* 100931	100931: contig of 9036 bp in length
* 100932	101031: gap of unknown length
* 101032	102653: contig of 1622 bp in length.
FEATURES	Location/Qualifiers
SOURCE	1..102653
	/organism="Lotus corniculatus var. japonicus"
	/mol_type="genomic DNA"
	/variety="japonicus"
	/db_xref="taxon:34305"
	/chromosome="3"
	/clone="LjT37F23"
	/clone_1b="LjT 1library"
	/note="TAC clone:TW0506, synonym:Lotus japonicus"
gap	772..871
	/estimated_length=unknown
gap	1868..1967
	/estimated_length=unknown
gap	2909..3008
	/estimated_length=unknown
gap	3973..4072
	/estimated_length=unknown
gap	5829..5928
	/estimated_length=unknown
gap	7171..7270
	/estimated_length=unknown
gap	8999..9098
	/estimated_length=unknown
gap	10096..10195
	/estimated_length=unknown
gap	11820..11919
	/estimated_length=unknown
gap	13526..13625
	/estimated_length=unknown
gap	15096..15195
	/estimated_length=unknown
gap	17492..17591
	/estimated_length=unknown
gap	19659..19758
	/estimated_length=unknown
gap	21398..21497
	/estimated_length=unknown
gap	23876..23975
	/estimated_length=unknown
gap	25653..25752
	/estimated_length=unknown
gap	26598..26697
	/estimated_length=unknown
gap	28741..28840
	/estimated_length=unknown
gap	31618..31717
	/estimated_length=unknown
gap	33601..33700
	/estimated_length=unknown
gap	36668..36767
	/estimated_length=unknown
gap	39713..39812
	/estimated_length=unknown
gap	42490..42589
	/estimated_length=unknown
gap	45917..46016
	/estimated_length=unknown
gap	49064..49163
	/estimated_length=unknown
gap	52245..52344
	/estimated_length=unknown
gap	55908..56007
	/estimated_length=unknown

gap	59275..59374	
	/estimated_length=unknown	
gap	61820..61919	
	/estimated_length=unknown	
gap	66982..67081	
	/estimated_length=unknown	
gap	72121..72220	
	/estimated_length=unknown	
gap	75655..75754	
	/estimated_length=unknown	
gap	81138..81237	
	/estimated_length=unknown	
gap	85716..85815	
	/estimated_length=unknown	
gap	91796..91895	
	/estimated_length=unknown	
gap	100932..101031	
	/estimated_length=unknown	
ORIGIN		
Query Match	12.9%; Score 103.8; DB 14; Length	
Best Local Similarity	53.3%; Pred. No. 4,66-20;	
Matches	281; Conservative 0; Mismatches 222; Indels	
OY	174	233
DB	14335	14276
OY	234	293
DB	14275	14216
OY	294	353
DB	14215	14156
OY	354	413
DB	14155	14105
OY	414	473
DB	14104	14045
OY	474	533
DB	14044	14000
OY	534	593
DB	13999	13940
OY	594	653
DB	13939	13880
OY	654	
DB	13879	
RESULT 7		
AF139499		
LOCUS	AF139499	1031 bp mRNA linear
DEFINITION	Prunus armeniaca p85RF mRNA, complete cds.	1999
ACCESSION	AF139499	
VERSION	AF139499.1 GI:5031280	
KEYWORDS		
SOURCE		
ORGANISM	Prunus armeniaca (apricot)	
	Prunus armeniaca	
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta	ta;
	Spermatophyta; Magnoliophyta; eudicotyledons; core	nb;
	rosidae; euroside I; Rosales; Rosaceae; Amygdaloidae	

REFERENCE 1 (bases 1 to 1031)
AUTHORS Mbegwe-A-Mbegwe,D. and File-Lycaon,B.R.
TITLE Molecular cloning and nucleotide sequence of a putative protein
JOURNAL from apricot (Prunus armeniaca var. Bergeron)
Unpublished
REFERENCE 2 (bases 1 to 1031)
AUTHORS Mbegwe-A-Mbegwe,D. and File-Lycaon,B.R.
TITLE Direct Submission
JOURNAL Submitted (31-MAR-1999) INRA, Domaine de Duclos, Petit-Bourg 97170,
France
FEATURES
source Location/Qualifiers
1..1031
/organism="Prunus armeniaca"
/mol_type="mRNA"
/varietal="Bergeron"
/db_xref="taxon:36596"
/cissue_type="Fruit mesocarp and endocarp"
/dev_stage="ripe"
1..1031
/gene="p85RP"
113..634
/gene="p85RP"
/note="similar to Phaseolis vulgaris PVPR3 protein"
/codon_start=1
/product="unknown"
/protein_id="ABD38146.1"
/db_xref="GI:5031281"
/translation="MEHERTGQPHPEBPILCVNNCGFSGSVATRNMGSKCHDKMLK
SEDAKLAASFENIVNGTNSNGNEPVVAADVDAHVEPKTISLOPSFSGSGG
SGEARPEGRKCGTCNKRGVLTGFCRCHLFCVHRHRSKDKCPYDHTAADIYAK
ANPVKADKLEKI"
ORIGIN
Query Match 12.7%; Score 102.2; DB 15; Length 1031;
Best Local Similarity 67.8%; Pred. No. 1.2e-19;
Matches 143; Conservative 0; Mismatches 68; Indels 0; Gaps 0;
QY 509 CCCAATCCATGGGGAATCTTTCAGAGAAAGCGTGTGATTCAGAGATTTAAGTGTGGCTGT 568
DB 443 CCAAAACGTTGGGGAATCTTTCAGAGAAAGCGTGTGATTCAGAGATTTAAGTGTGGCTGT 502
QY 569 GGAACCTCTACTGAGCTTATCATCGGTACTCGAGCAAAACATTCGACATATGACTAC 628
DB 503 GGTACCTTTTCTGTCAGTATCATGCTTATTCAGAGAAACATGACTGCCCTTATGATTAAC 562
QY 629 AAAGCCGAGGAGGAGAGAGCGATTGCGAAAGCTTAATCTTGTGTGCGCCGAGAGGTT 688
DB 563 CACACTGCTGACGCGAGATGATGCTTAAGCCAAACCGTGTGTAAGGCTGATTAAGCTT 622
QY 689 GTCAAGTTTATGATGAGCATCCGTTAAGCTTT 719
DB 623 GAAAAAATCTAAGCTGATGTGTGAAGTTT 653
RESULT 8
LOCUS BT016379 1020 bp mRNA linear PLN 27-OCT-2004
DEFINITION Zea mays clone Contig212 mRNA sequence.
ACCESSION BT016379
VERSION BT016379.1 GI:54651160
KEYWORDS FLI CDNA.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 1020)
La1,J., Dey,N., Kim,C.-S., Bharti,A.K., Rudd,S., Mayer,K.F.X.,
Larkin,B., Becraft,P. and Messing,J.
Characterization of the maize endosperm transcriptome and its
comparison to the rice genome
Unpublished
JOURNAL 2 (bases 1 to 1020)
REFERENCE

AUTHORS La1,J., Dey,N., Kim,C.-S., Bharti,A.K., Rudd,S., Ma
Larkin,B., Becraft,P. and Messing,J..
TITLE Direct Submission
JOURNAL Submitted (27-OCT-2004) Waksman Institute, Rutgers
Frelinghuysen Rd, Piscataway, NJ 08854, USA
FEATURES
source Location/Qualifiers
1..1020
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="taxon:4577"
/clone="Contig212"
ORIGIN
Query Match 12.7%; Score 101.8; DB 15; Length
Best Local Similarity 53.5%; Pred. No. 1.6e-19;
Matches 301; Conservative 0; Mismatches 232; Indels
QY 156 AGGAGACGACCTCGGAGGCCCCGTGAGGGTCCAGTTATTCGAAAGACCT 215
DB 117 AGGAGGCTGGGTGCGAGGCCCCCGAGGAGCCCATCTCTGCACTCAATTA 176
QY 216 TCGGAGCCCAAGCTACATGGGGTGTGCTCGAAGTGTACCGAGAGAC 275
DB 177 TCGGAGCGCGCGGACATGAACATGTCTTAAGTGCACAGAGAGAT 236
QY 276 CGAAGATACGCGCTTAACTGACAGGCAACTCAGGCTGCTCAGGCGAC 335
DB 237 AGGATCAGGCCCAAGCTGCTGC-----CTCTCTATACAC 282
QY 336 CTGCTGCTGTTACGCCCCCGCTCTCTATACATGAGACCAAGCTCACAT 395
DB 283 CGGACGACCGCGT-----CATGAGCCGGTTGTGCT 328
QY 396 GAACATGATTTGCGCGCATCATCTTCAGCTATCAACAGACCTGGT 455
DB 329 GTAGTAGCTGTGCGGCAAGTGAAGTGAACAAATGAAC--GTGACGCA 386
QY 456 CAGCTGCCCTCAGGAGGATGAATCTCTTATGCGAGCTCCCTTAAGAC 515
DB 387 TTGCGGAGCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 446
QY 516 GATGCGGATCTTTCAGAGAAAGGATGTTGATTCAGAGATTTAAGTGTG 575
DB 447 GGTGACGCGCTTCAGAGAAAGGATGTTGAGCTCAGCGGATTCACTGCGCC 506
QY 576 TCTACTGCGCTTTCATGCTGTAATCTGCAAAACACATCTTGACATATGA 635
DB 507 TGTACTGCGCGCTCCACCGCTACTCCGACAGCAGCATGCAAGTTGCA 566
QY 636 CAGGCGAGAAAGCATTCGAAAGCTTAATCTTGTGTCGCGCGAGAA 695
DB 567 CTGCGAGGAGCGCCATTCGAAAGCTTAATCCGAGTGTGAAGGACAGACAN 626
QY 696 TTGATGAGCATCCGTTAAGCTT 718
DB 627 TCTAGGGGGGATCCCTTAAGGTT 649
RESULT 9
LOCUS AY15607 1162 bp mRNA linear -2004
DEFINITION Zea mays putative zinc finger protein Zmzf mRNA, cc
ACCESSION AY15607
VERSION AY15607.1 GI:41350258
KEYWORDS
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; i
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 1162)
Zhang,Z.-X., Tang,W.-H., Tao,Y.-S. and Zheng,Y.-L.
cDNA microarray analysis of early response in maize
Unpublished
JOURNAL
REFERENCE

roots to submerging stress

JOURNAL

Unpublished

2 (bases 1 to 1162)

Zhang,Z.-X., Tang,W.-H., Tao,Y.-S. and Zheng,Y.-L.

AUTHORS

TITLE

Submitted (01-JAN-2004) Maize Group of National Key Laboratory of

Crop Genetic Improvement, Huazhong Agricultural University of

China, Wuhan, Hubei 430070, China

FEATURES

source

1..1162

/organism="Zea mays"

/mol_type="mRNA"

/culivar="m017"

/db_xref="taxon:4577"

/note="authority: Zea mays L."

106..807

/codon_start=1

/product="putative zinc finger protein 2m2f"

/protein_id="AA00453.1"

/db_xref="GI:41350259"

/translation="MVSPPDLKIKWGRGASASRSPSPPLLYLIGFPASFGSS

ERNRKGRLIRKEMEKKEACQAPBEPILICINNGFFGSATNMKSKHEKITK

ODQALAASSIDSLVDGSDAVMEPGAGSNTVVAQVELKKNVQPADVAEPVEG

AAISKGVGPMRCASCRKRVGLTGFNCRGNTLYCALHRYSPKDKEDRYTAADAI

AKANPVVADKLDKI"

ORIGIN

Query Match 12.5%; Score 100.6; DB 15; Length 1162;

Best Local Similarity 52.2%; Pred. No. 3,8e-19;

Matches 294; Conservative 0; Mismatches 239; Indels 30; Gaps 2;

156 AGGAGACGACCTCGAGAGCCCTGAGGGGTCCAGTTATGTGCAAGAACCTTTGGGCTCT 215

293 AGGAGGCTGGGTGCGAGGCCCGCCGAGAGGCCATCTCTGCAATCAATTACTTGGCTTCT 352

216 TCGGAGCCGACGACATCCATGAGGCTTGTGCTCGAAGTGTCTACGAGACAGTCATGCAAG 275

353 TCGGAGCGCGCGGACCATGAACATGTGCTCCAGAGTGCACAAAGAGATGATTAAGAAAGC 412

216 CGAAGATGACGGCTTTAGCTGAGCAAGCCACTCAGGCTGCTTGAAGGACATGTGCACAG 335

413 AGGATCAGGCCCAAGCTGGCTGC-----CTCCTCTATCGACAGCATGTGGA 458

336 CTGCTGCTGTTCAGCCCCCGCTCCTGTATCATGAGACCAAGCTCAATGCGAGTTGAGA 395

459 CGGACGCGACGCTGATAGAGCCCGGTGTGCTGCAAGAACACGATGATGCTGTTGC 518

396 GAACAATGATTGTGCGCATCAATCTTCCAGCTATCAACAAGACCTGTGTTACCCCGCTG 455

519 CCAAGTTGATTGCAAAAAGACGTGAGC-----AGCCGCTGATG 562

456 CAGCTGCCCTCAGGACGTAAGTCTCTATTCGACAGTCCCTCTGAGCCCGAGCCCAATC 515

553 TTGCCGAACCCCAACAGAGGGGTGTGCGCGCATCTCCAAAGGGGGAAGGTAGGGCGGAAC 622

516 GATGCGGATCTTTCGAGAAAGCGTGTGATGACAGGATTTAAGTGTGCTGTGCAACC 575

623 GGTGACGCGCTTCGAGAAAGGTTGACTTAACGAGATTCACTGCGGTGTGGAACCT 682

683 GTGTATGCGCACTCCACCGCTACTCGACAAACAGACGATGCAAGTTCCACTACCGGACG 742

636 CAGGCGAGAGCGATTGTGGAAGTAACTCTCTTGTCTGTGCGGCGGAGAAAGTTGTCAAGT 695

743 CTGCCAGGAGACCCATTGCAAGGCTAATCCGATGTGTAAGGACAGCAAGCTCGACAAGA 802

696 TTTGATGAGCATCCGTTAAGCTT 718

803 TCTAGGGGGGTTCCCTACGTT 825

ReguLr 10

BT017601

LOCUS BT017601 941 bp mRNA linear

DEFINITION Zea mays clone EL01N0431H03.c mRNA sequence.

ACCESSION

BT017601

VERSION BT017601.1 GI:54652382

KEYWORDS

FLI CDNA.

SOURCE

Zea mays

ORGANISM

Eukaryote; Viridiplantae; Streptophyta; Embryophyta

Spermatophyta; Magnoliophyta; Liliopsida; Poales; P

Clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 941)

Lat,J., Dey,N., Kim,C.-S., Bharti,A.K., Rudd,S., Ma

Larking,B., Berafai,P. and Messing,J.

Characterization of the maize endosperm transcript

comparison to the rice genome

Unpublished

JOURNAL

2 (bases 1 to 941)

Lat,J., Dey,N., Kim,C.-S., Bharti,A.K., Rudd,S., Ma

Larking,B., Berafai,P. and Messing,J.

Submitted (27-OCT-2004) Waksman Institute, Rutgers

Freelightinghway Rd, Piscataway, NJ 08854, USA

FEATURES

source

1..941

/organism="Zea mays"

/mol_type="mRNA"

/db_xref="taxon:4577"

/clone="EL01N0431H03.c"

ORIGIN

Query Match 12.1%; Score 97; DB 15; Length 941

Best Local Similarity 52.9%; Pred. No. 4,9e-18;

Matches 298; Conservative 0; Mismatches 235; Indels 3;

156 AGGAGACGACCTCGAGAGCCCTGAGGGTCCAGTTATGTGCAAGAACCT 215

123 AGGAGGCTGGGTGCGAGGCCCGCCGAGGAGCCATCTCTGATCAGGGG 182

216 TCGGAGCCGACGACATCCATGAGGCTTGTGCTCGAAGTGTCTACCGAGAGC 275

183 TCGGAGCGCGCGACCATGAACTGTCTTAAGTGTGCAAGAGAT 242

276 CGAAGATGACGGCTTTAGCTGAGCAAGCCACTCAGGCTGCTCAGGCGAC 335

243 AGGATCAGGCCCAAGCTGCTC-----CTCCTCTATCGAC 288

336 CTGCTGCTGTTCAGCCCCCGCTCCTGTATCATGAGACCAAGCTCAATG 395

289 CGGACGAGAGCCCGT-----CATGAGCCGGTTGTTGCTG 334

396 GAACAATGATTGTGCGCATCAATCTTCCAGCTATCAACAAGACTGTGT 455

335 GAGATGCTGTGCGCAAGTGAATTCGAGTTGCAACATGAAC--GTGACAGA 392

456 CAGTGCCTTCAGGACGTAAGTCTCTATTCGACAGTCCCTTAGACC 515

393 TTGCCGACCCAGAGAGGGGGGGGGGAGATCTCCAAAGGGGGAAGG 452

516 GATGCGGATCTTTCGAGAAAGCGTGTGATGACAGATTTAAGTGTG 575

453 GGTGACGCGCTTCGAGAAAGGTTGGGCTCACGGGATTCACATGCGG 512

576 TCTACTGCGCTTTACATGGGTACTCGACAAACACACTTGCACATATGA 635

513 GTTACTGGGCGCTCCACCGGTACTCCGACAAAGCAAGCTGCAAGTTCA 572

636 CAGGCGAGAGCGATTGCGAAAGCTAATCTCTTGTGCTGCGCGAGAA 695

573 CTGCCAGGAGCGCATTTGCAAGGCTAATCCGGTGTGTAAGGACAGACA 632

696 TTTGATGAGCATCCGTTAAGCTT 718

Db	633	CTGAGGGGGGTTCCCTACGGT	655
RESULT 11			
LOCUS	AR252087	283 bp	DNA
DEFINITION	Sequence 7446 from patent US 6476212.		linear
ACCESSION	AR252087		PAT 20-DEC-2002
VERSION	AR252087.1	GI:27299961	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 283)		
AUTHORS	Laigudi, R.V., Ito, L.Y. and Sherman, B.K.		
TITLE	Polynucleotides and polypeptides derived from corn ear		
JOURNAL	Patent: US 6476212-A 7446 05-NOV-2002;		
FEATURES	Incyte Genomics, Inc.; Palo Alto, CA		
SOURCE	location/Qualifiers		
	1..283		
	/organism="unknown"		
	/mol_type="genomic DNA"		
ORIGIN			
Query Match	12.0%; Score 96.8; DB 6; Length 283;		
Beet local similarity	67.2%; Pred. No. 5,4e-18;		
Matches 137; Conservative	0; Mismatches 67; Indels 0; Gaps 0;		
QY	509	CCCAATCGATCGGATCTTGGCAGAAAGCGTGTGGATTGACAGATTTAAGTGGCTGT	568
Db	2	CCAAACGGGTGGCAACCTGTATGGAAGCGTGTGGGTTGACGGGTTTAACTGCCGATC	61
QY	569	GGCAACCTCTACTGCGCTTTACATCGTACTCGCAACAACACACTTGCATATGACTAC	628
Db	62	GGGAACAGCTACTGTGTGATGACCGGCTACTCCGAACAACAGATGCGAGTTGCACAT	121
QY	629	AAAGCCGAGGCGCAGAAAGCGATTGCCAAAGCTAATCTCTTGTCTGTCGCGAGAGGTT	688
Db	122	CGAATCGAGGTAGGGGAGCGCTATCGCAAGGCGCATTCAGTGTGAGACGGAGAGGCTT	181
QY	689	GTCAGTTTGTGATGAGCATCCGCTT	712
Db	182	GACAAATCTTAAGGCGGGGCGCATTT	205
RESULT 12			
LOCUS	BC059673		
DEFINITION	BC059673	2022 bp	mRNA linear
ACCESSION	BC059673		VRT 20-OCT-2004
VERSION	BC059673.1	GI:37589767	
KEYWORDS			
SOURCE	Danio rerio (zebrafish)		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes; Cyprinidae; Danio.		
REFERENCE	1 (bases 1 to 2022)		
AUTHORS	Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, J., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heish, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Cesaevant, T.L., Scheetz, T.E., Brownstein, M.J., Udutin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Rana, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullen, S.J., Bosak, S.A., McEwan, P.J., McEwan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Woley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Pahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,		

[illegible]

QY 540 TTGATTCAGAGATTTTAAGTTCGCTGCGCAACCTCTACTGCGCTTTACATCGGTA 599
DB 610 TCGGCTTAACAGAGTTTGAATCTGATGCTGTAATCTGTTCTGTGGAATTCATCGGTA 669
QY 600 CGGACAAACACACTTTCACATATGACTACAAAGCCGAGGAGGAGAGCGATTTGCGAA 659
DB 670 CAGACAGACACAACTGACATACATATGATTAAGAGGCGAAAGCCGCCCAAGATCCGCAAG 729
QY 660 CTATCTCTCTGCTGCGCCGAGAA 684
DB 730 AAATCTCTGATGATGCGCGATTA 754

RESULT 13
LOCUS AR246592 281 bp DNA linear PAT 20-DEC-2002
DEFINITION Sequence 1951 from patent US 6476212.
ACCESSION AR246592
VERSION AR246592.1 GI:27294466
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE
AUTHORS 1 (bases 1 to 281)
TITLE laijudi, R.V., Ito, L.Y. and Sherman, B.K.
JOURNAL Polynucleotides and polypeptides derived from corn ear
Patent: US 6476212-A 1951-05-NOV-2002;
Incyte Genomics, Inc.; Palo Alto, CA
FEATURES
source
1..281
/organism="Unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 11.6%; Score 93.6; DB 6; Length 281;
Best Local Similarity 67.3%; Pred. No. 5, 2e-17;
Matches 132; Conservative 0; Mismatches 64; Indels 0; Gaps 0;
QY 509 CCCATTCAGATGGGATCTTTCAGAGAAAGCGTGTGATTCAGAGATTTAAAGTGTGCTGT 568
DB 10 CCAGACCGGTGCGCGCGCTGCAAGAGCGTGTGGCTGACCGGATTTAGCTGCCGATGC 69
QY 569 GCGAACCCTTACTGCGCTTTTACATCGTACTCGACAAACACACTTTCACATATGACTAC 628
DB 70 GGGAACTATGATCTGTTGCGTGCACCGCTACTCGACAAACAGACTGTGATTCGACTAT 129
QY 639 AAAGCCGAGGCGAGAGAGCGATTTGCGAAAGCTAATCTCTTGTGTGCGCGGAGAGTT 688
DB 130 CGACTGCGAGCAAGGAGCGGATTCGCAAGGCCAATCTGTGTGTGAGGCGCGAGAGCTC 189
QY 689 GTCAAGTTTGTATGAG 704
DB 190 GACAAAGATCTGAGCG 205

RESULT 14
LOCUS BT019326 779 bp mRNA linear PLN 27-OCT-2004
DEFINITION Zea mays clone Contig1045.F mRNA sequence.
ACCESSION BT019326
VERSION BT019326.1 GI:54654107
KEYWORDS FLI CDNA.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
AUTHORS lai, J., Dey, N., Kim, C.-S., Bhatti, A.K., Rudd, S., Mayer, K.F.X., Larkins, B., Bectraft, P. and Messing, J.
TITLE Characterization of the maize endosperm transcriptome and its comparison to the rice genome
JOURNAL Unpublished

REFERENCE 2 (bases 1 to 779)
AUTHORS lai, J., Dey, N., Kim, C.-S., Bhatti, A.K., Rudd, S., Ma, Larkins, B., Bectraft, P. and Messing, J.
TITLE Direct Submission
JOURNAL Submitted (27-OCT-2004) Waksman Institute, Rutgers
Frelinghuysen Rd, Piscataway, NJ 08854, USA
FEATURES
source
1..779
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="taxon:4577"
/clone="Contig1045.F"

ORIGIN
Query Match 11.6%; Score 93.6; DB 15; Length 7
Best Local Similarity 67.3%; Pred. No. 5.4e-17;
Matches 132; Conservative 0; Mismatches 64; Indels 0;
QY 509 CCCATTCAGATCGGATCTTTCAGAGAAAGCGTGTGATTCAGAGATTTTA 568
DB 102 CCAGACCGGTGCGCGCGCTGCAAGAGCGTGTGGCTGACGGAATTTA 161
QY 569 GCGAACCCTTACTGCGCTTTTACATCGTACTCGACAAACACACTTGA 628
DB 162 GGGAACTATGATCTGTTGCGTGCACCGCTACTCGACAAACAGACTGTC 221
QY 629 AAAGCCGAGGCGAGAGAGCGATTTGCGAAAGCTAATCTCTTGTGCTG 688
DB 222 CGACTGCGAGCAAGGAGCGGATTCGCAAGGCCAATCTGTGTGAGCG 281
QY 689 GTCAAGTTTGTATGAG 704
DB 282 GACAAAGATCTGAGCG 297

RESULT 15
LOCUS BT014337 1026 bp mRNA linear
DEFINITION Lycopersicon esculentum clone 133610F, mRNA sequenc
ACCESSION BT014337
VERSION BT014337.1 GI:47105752
KEYWORDS FLI CDNA.
SOURCE Lycopersicon esculentum (Solanum lycopersicum)
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta
Spermatophyta; Magnoliophyta; eudicotyledons; core
asterids; lamids; Solanales; Solanales; Solanum; L
REFERENCE
AUTHORS Kirkness, E.F., Wang, W. and Vazelli, A.
TITLE 1 (bases 1 to 1026)
JOURNAL Submitted (11-MAY-2004) The Institute for Genomic R
Medical Center Drive, Rockville, MD 20850, USA
FEATURES
source
1..1026
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/db_xref="taxon:4081"
/clone="133610F"
/tissue_type="fruit"
/note="TX1178"

ORIGIN
Query Match 11.6%; Score 93; DB 15; Length 102
Best Local Similarity 67.0%; Pred. No. 8.4e-17;
Matches 132; Conservative 0; Mismatches 65; Indels 0;
QY 506 GAGCCCATTCAGATCGGATCTTTCAGAGAAAGCGTGTGATTCAGAGAT 565
DB 336 GGGCCAAACAGATGACGACCTGTAGAGAGAGAGATGTTGACCGGCT 395
QY 566 TGTGGCAACTCTTACTGCGCTTTTACATCGGATTCGAGCAACACACTT 625
DB 396 TGTGGAAATCTTCTCTGCTCAGCTCATGCTACTACAGCAACATGAGCT 455

Qy 626 TACAAAGCCGACAGGAGAGCGATTGCGAAAGCTATCTCTTGTCTGCGCCGAGAG 685
Db 456 TATCGAAGGCTGCTCAAGATGCTATTTGCGAAAGCCACCAATTGTTAAGGCTGAGAAAG 515
Qy 686 GTTGTCAAGTTTGATG 702
Db 516 CTTGACAAAATATGAG 532

Search completed: December 8, 2005, 09:09:44
Job time : 4261 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 8, 2005, 06:02:18 ; Search time 552 Seconds
(without alignment)
9707.270 Million cell updates/sec

Title: US-10-716-089-12

Perfect score: 804
Sequence: 1 atcccgagccaccagtcgccgc.....agtcgcatcaagagctcgcc 804

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : N_Geneseq_21:*
- 1: geneeqn1980s:*
 - 2: geneeqn1990s:*
 - 3: geneeqn2000s:*
 - 4: geneeqn2001as:*
 - 5: geneeqn2001bs:*
 - 6: geneeqn2002as:*
 - 7: geneeqn2002bs:*
 - 8: geneeqn2003as:*
 - 9: geneeqn2003bs:*
 - 10: geneeqn2003cs:*
 - 11: geneeqn2003ds:*
 - 12: geneeqn2004as:*
 - 13: geneeqn2004bs:*
 - 14: geneeqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	804	100.0	804	6	ABST0622 Transcrip
2	804	100.0	804	6	AA955832 CDNA enco
3	776.4	96.6	1531	6	ABST0614 Transcrip
4	776.4	96.6	1531	6	AA955824 Partial c
5	103.4	12.9	1003	13	ADK47821 Plant ful
6	103	12.8	873	13	ADK53121 Plant ful
7	103	12.8	945	13	ADK51406 Plant ful
8	102.6	12.8	932	13	ADK51034 Plant ful
9	102.2	12.7	1032	13	ADK52834 Plant ful
10	101.8	12.7	776	13	ADK47307 Plant ful
11	101.8	12.7	821	13	ADK47342 Plant ful
12	101.8	12.7	821	13	ADK47342 Plant ful
13	101.8	12.7	826	13	ADK46752 Plant ful
14	101.8	12.7	841	13	ADK50951 Plant ful
15	101.8	12.7	846	13	ADK50987 Plant ful
16	101.8	12.7	853	13	ADK54045 Plant ful
17	101.8	12.7	863	13	ADK52432 Plant ful
18	101.8	12.7	863	13	ADK53800 Plant ful
19	101.8	12.7	868	13	ADK33018 Plant ful

20	101.8	12.7	868	13	ADK54028	Adx	ful
21	101.8	12.7	872	13	ADK51702	Adx	ful
22	101.8	12.7	876	13	ADK52889	Adx	ful
23	101.8	12.7	886	13	ADK12369	Adx	ful
24	101.8	12.7	894	13	ADK48800	Adx	ful
25	101.8	12.7	898	13	ADK61184	Adx	ful
26	101.8	12.7	899	13	ADK50721	Adx	ful
27	101.8	12.7	902	13	ADK52776	Adx	ful
28	101.8	12.7	904	13	ADK47138	Adx	ful
29	101.8	12.7	918	13	ADK61202	Adx	ful
30	101.8	12.7	920	13	ADK47807	Adx	ful
31	101.8	12.7	924	13	ADK60812	Adx	ful
32	101.8	12.7	926	13	ADK52877	Adx	ful
33	101.8	12.7	928	13	ADK52869	Adx	ful
34	101.8	12.7	942	13	ADK55778	Adx	ful
35	101.8	12.7	943	13	ADK5613	Adx	ful
36	101.8	12.7	948	13	ADK10810	Adx	ful
37	101.8	12.7	953	13	ADK53073	Adx	ful
38	101.8	12.7	956	13	ADK53049	Adx	ful
39	101.8	12.7	959	13	ADK47752	Adx	ful
40	101.8	12.7	976	13	ADK47829	Adx	ful
41	101.8	12.7	978	13	ADK32914	Adx	ful
42	101.8	12.7	985	13	ADK47720	Adx	ful
43	101.8	12.7	987	13	ADK62462	Adx	ful
44	101.8	12.7	989	13	ADK48942	Adx	ful
45	101.8	12.7	990	13	ADK46608	Adx	ful

ALIGNMENTS

RESULT 1
ABST0622 ID ABST0622 standard; CDNA, 804 BP.
XX
AC ABST0622;
XX
DT 28-NOV-2002 (first entry)
XX
DE Transcription factor stress-related protein (TFSRP), cDNA
XX
KW Transcription factor stress-related protein; TFSRP; MYB-1;
KW AP2 similar-2 protein; APS-2; zinc-finger factor-2; ZF-2;
KW ZF-4 protein; ZF-5 protein; CANT-Box binding factor protei;
KW Sigma factor like protein; SFU-1; transgenic; plant; droug;
KW environmental stress; Physcomitrella patens; high salinity
KW ciliate; fungus; gene; ss.
XX
OS Physcomitrella patens.
XX
PN US2002102695-A1.
XX
PD 01-AUG-2002.
XX
PF 06-APR-2001; 2001US-00828303.
XX
PR 07-APR-2000; 2000US-0196001P.
XX
PA (SILV/) SILVA O D C E.
PA (BOHN/) BOHNERT H J.
PA (THIELE/) THIELEEN N V.
PA (CHEN/) CHEN R.
PI Silva O DCE, Bohnert HJ, Thieleen NV, Chen R;
DR WPI: 2002-690614/74.
DR P-PSDB; ABG93809.
XX
PT Novel transcription factor stress-related protein such as
PT protein, zinc-finger protein, or CANT-Box binding factor i;
XX useful for modifying stress tolerance of plant.
XX
PS Claim 3; Fig 2D; 104pp; English.

xx The invention relates to an isolated transcription factor stress-related protein (TFSRP) (I) selected from AP2 similar-2 protein (AP2-2), zinc-finger factor-2 (ZF-2) protein, ZF-3 protein, ZF-4 protein, ZF-5 protein, MYB-1 protein, CMT-Box binding factor protein-3 (CMT-3), Sigma factor like protein (SFL-1) and their orthologues. Also described is: (1) an isolated TFSRP coding nucleic acid (II) coding for (I); (2) a transgenic plant cell (III) transformed by (II), where the expression of (II) in the plant cell results in increased tolerance to an environmental stress as compared to a wild-type variety of the plant cell; (3) a transgenic plant (IV) comprising (III); (4) a seed (V) produced by (IV), where the seed is true breeding for an increased tolerance to environmental stress as compared to a wild-type variety of the plant cell; (5) an agricultural product (VI) produced by (IV) or (V); and (6) an isolated recombinant expression vector (VII) comprising (II), where expression of (VII) in a host cell results in increased tolerance to environmental stress as compared to a wild-type variety of the host cell. (II) is useful for modulating (increasing or decreasing) stress tolerance of a plant (transgenic or not transgenic), by modifying the expression of TFSRP in the plant. The plant is transformed with a promoter that directs expression of the TFSRP. The promoter is tissue specific and is developmentally regulated. TFSRP expression is modified by administration of an antisense molecule that inhibits expression of TFSRP. (VII) is useful for producing a transgenic plant containing (II), where expression of the nucleic acid in the plant results in increased tolerance to environmental stress as compared to a wild-type variety of the plant, by transforming a plant cell with (VII) comprising the nucleic acid, generating from the plant cell a transgenic plant with an increased tolerance to environmental stress as compared to a wild-type variety of the plant. (II) is useful for conferring stress tolerance such as drought, cold and/or salt tolerance to plants. (I), (II), (III) or (VII) is useful for identifying Physcomitrella patens and related organisms, for mapping of genomes of organisms related to P. patens, for identifying and localising P. patens sequences of interest, for evolutionary studies, for determining TFSRP regions required for function, for modulating TFSRP activity, for modulating metabolism of one or more compounds, and for modulating stress resistance. (I) is useful for reducing stress tolerance response plants or more particularly, in the transcription of a protein involved in a stress tolerance response in a P. patens plant. (II) is useful for transforming plants and thus inducing tolerance to stresses such as drought, high salinity and cold, for identifying the presence of P. patens or a related organism in a mixed population of microorganism, serve as marker for specific regions of the genome for mapping the genome and for the functional studies of P. patens proteins. (II) is useful as reference points for mapping the moss genomes, or of genomes of related organisms, for evolutionary and protein structural studies, for generating knockout mutation in the genomes of various organisms such as bacteria, mammalian cells, yeast cells and plant cells, useful for evaluating their ability or capacity to tolerate various stress conditions and the effect on the phenotype and/or genotype of the mutation. (II) is useful as marker for the construction of a genomic map in related mosses. (I) or (II) is useful for generating algae, ciliates, plants, fungi or other microorganisms expressing mutated TFSRP nucleic acid and protein molecules for improving stress tolerance. ABS70611-ABS70681 represent P. patens TFSRP coding sequences and PCR primers of the invention.

xx Sequence 804 BP, 186 A; 214 C; 205 G; 199 T; 0 U; 0 Other;

xx

Query Match 100.0%; Score 804; DB 6; Length 804;
Best Local Similarity 100.0%; Pred. No. 2.3e-250;
Matches 804; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATCCGGGACCAAGTCCTTATGTCATTAAGTGTGGTTCGAAGTCTGAAGC 60
|||
DB 1 ATCCGGGACCAAGTCCTTATGTCATTAAGTGTGGTTCGAAGTCTGAAGC 60
|||
QY 61 CTGAGCGAGATTTGCGAGATTTTCTCATAGCCTTCTGATTTGGAAGATACCTTAT 120
|||
DB 61 CTGAGCGAGATTTGCGAGATTTTCTCATAGCCTTCTGATTTGGAAGATACCTTAT 120
|||
QY 121 TAGTCTGTTAAAGATGCGACCGAGCGTGTCTCAGAGACGACCTCGAGCCCTGA 180
|||

DB 121 TAGTCTGTTAAAGATGCGACCGAGCGTGTCTCAGAGACGACCTCG 180
|||
QY 181 GGGTCAGATTATGTCGAAGAACCCTTGGCGCTTCTTGGGCGCCAGGT 240
|||
DB 181 GGGTCAGATTATGTCGAAGAACCCTTGGCGCTTCTTGGGCGCCAGGT 240
|||
QY 241 GTGCTCGAAGTGTACCGAGAGACAGTCATGCAAGCGAAGATGACGGCT 300
|||
DB 241 GTGCTCGAAGTGTACCGAGAGACAGTCATGCAAGCGAAGATGACGGCT 300
|||
QY 301 AGCCACTGAGGCTCTCAGCGACATCTGCCACAGCTGCTGCTTACG: 360
|||
DB 301 AGCCACTGAGGCTCTCAGCGACATCTGCCACAGCTGCTGCTTACG: 360
|||
QY 361 TGTACATGAGCCAAAGCTCAATCGAGGTTGAGAGAACAATGATTTGTC 420
|||
DB 361 TGTACATGAGCCAAAGCTCAATCGAGGTTGAGAGAACAATGATTTGTC 420
|||
QY 421 TTCCAGCTATCAACAAGACCTGTTACCCCGCTGCAGCTGCCCTCAG: 480
|||
DB 421 TTCCAGCTATCAACAAGACCTGTTACCCCGCTGCAGCTGCCCTCAG: 480
|||
QY 481 CTCTATGCGACGCTCCTTACACCCGAGCCCAATCGATGCGGATCTTGC 540
|||
DB 481 CTCTATGCGACGCTCCTTACACCCGAGCCCAATCGATGCGGATCTTGC 540
|||
QY 541 TGGATTGACAGATTAAAGTGTGCTGTGGCAACCTCTACGCGCTTTA 600
|||
DB 541 TGGATTGACAGATTAAAGTGTGCTGTGGCAACCTCTACGCGCTTTA 600
|||
QY 601 GGACAAACACACTTGCACATATGACTACAAAGCCGACGAGAGAGG 660
|||
DB 601 GGACAAACACACTTGCACATATGACTACAAAGCCGACGAGAGAGG 660
|||
QY 661 TAACTCTTGTGCTGGCGGAGAAAGTTGTCAAGTTTGTATGACATTC 720
|||
DB 661 TAACTCTTGTGCTGGCGGAGAAAGTTGTCAAGTTTGTATGACATTC 720
|||
QY 721 CTGCGCAGATTAGGCTTCTACATGAGTAACTTCAATCTTCTTC 780
|||
DB 721 CTGCGCAGATTAGGCTTCTACATGAGTAACTTCAATCTTCTTC 780
|||
QY 781 AGCGATGCGATCAAGAGCTCGCC 804
|||
DB 781 AGCGATGCGATCAAGAGCTCGCC 804
|||

RESULT 2
AAS95832
ID AAS95832 standard; cDNA; 804 BP.
AC AAS95832;
XX
AC 26-FEB-2002 (first entry)
XX
DE cDNA encoding zinc finger-4 (ZF-4).
XX
XX Transcription factor stress-related protein; AP2 similar-2
OS AP2-2; zinc-finger factor protein; ZF-2; ZF-3; ZF-4; ZF-5;
KM MYB-1; MYB-1 protein; CMT-box binding factor protein-3; CMT
KW Sigma factor like protein; SFL-1; environmental stress; mo.
XX 88.
XX Physcomitrella patens.
OS
XX
PN MO200177311-A2.
XX
PD 18-OCT-2001.
XX
PF 06-APR-2001; 2001WO-US011393.
XX
PR 07-APR-2000; 2000US-0196001P.

XX (BAD) BASF PLANT SCI GMBH.
 XX Da Costa Silva O, Bohnert HJ, Van Thielén N, Chen R;
 XX MPI; 2002-049146/06.
 XX P-PSDB; NAU1800.
 XX
 XX New polypeptide, useful for identification of Physcomitrella patens, and
 XX for modulating stress resistance of a plant, comprises an isolated
 XX transcription factor stress-related protein derived from Physcomitrella
 XX patens.
 XX
 XX Claim 14; Fig 2D; 129pp; English.
 XX
 XX The invention relates to an isolated transcription factor stress-related
 XX protein (TFSRP) (i) selected from an AP2 similar-2 protein (APS-2), a
 XX zinc-finger factor-2 protein (ZF-2), ZF-3 protein, ZF-4 protein, ZF-5
 XX protein, a MYB-1 protein (MYB-1), a CAAT-box binding factor protein-3
 XX (CABP-3), a Sigma factor like protein (SFL-1) and its orthologues. Also
 XX described is an isolated TFSRP coding nucleic acid (ii) which codes for
 XX (i); and an isolated recombinant expression vector (iii) comprising (ii),
 XX where expression of (iii) in a host cell results in increased tolerance
 XX to environmental stress as compared to a wild type variety of the host
 XX cell. (i) or (ii) is useful as markers for specific regions of the TFSRP
 XX genome. (i) or (ii) is useful for identification of Physcomitrella patens
 XX and related organisms, mapping of genomes of organisms related to P.
 XX patens, identification and localization of P. patens sequence of
 XX interest, evolutionary studies, determination of TFSRP regions required
 XX for function, modulation of a TFSRP activity, modulation of the
 XX metabolism of one or more cell functions, modulation of the transmembrane
 XX transport of one or more compounds and modulation of stress resistance.
 XX (iii) is useful for generating probes and primers for identifying and/or
 XX cloning TFSRP homologues in other cell types and other organisms, as well
 XX as TFSRP homologues from mosses and related species, and for evolutionary
 XX and protein structural studies. AAG59821-AAG59891 represent P. patens
 XX TFSRP coding sequences and PCR primers of the invention
 XX
 XX Sequence 804 BP; 186 A; 214 C; 205 G; 199 T; 0 U; 0 Other;
 XX
 XX Query Match 100.0%; Score 804; DB 6; Length 804;
 XX Best Local Similarity 100.0%; Pred. No. 2,3e-250;
 XX Matches 804; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX 1 ATCCGGGACCAAGCTCCGCTTATGCTGTGTCATTAGCTGTGTCAGCTTGAAGC 60
 XX 1 ATCCGGGACCAAGCTCCGCTTATGCTGTGTCATTAGCTGTGTCAGCTTGAAGC 60
 XX
 XX 61 CTGAGCGGATTTGAGAGATTTTCTCAACGCTTCTGATTAGAAAGTAAACCTTAT 120
 XX 61 CTGAGCGGATTTGAGAGATTTTCTCAACGCTTCTGATTAGAAAGTAAACCTTAT 120
 XX
 XX 121 TAGCTGTTAAAGATGCGCACCGAGCGTGTCTCAGAGACGACTCGACGCCCTGA 180
 XX 121 TAGCTGTTAAAGATGCGCACCGAGCGTGTCTCAGAGACGACTCGACGCCCTGA 180
 XX
 XX 181 GGGTCCAGTTATGTGCAAGAACCTTTGCGGCTTCTCGGACCAAGCTACCATGGGGTT 240
 XX 181 GGGTCCAGTTATGTGCAAGAACCTTTGCGGCTTCTCGGACCAAGCTACCATGGGGTT 240
 XX
 XX 241 GGTGCTGAAAGTGTCTCCGAGACAGTCAATGCAAGGAAGATGAGCGCTTGTAGCTGAGA 300
 XX 241 GGTGCTGAAAGTGTCTCCGAGACAGTCAATGCAAGGAAGATGAGCGCTTGTAGCTGAGA 300
 XX
 XX 301 AGCCACTGAGGCTGCTCAGGCGACATCTGCCACAGCTGTGCTTACGCCCCGCTCC 360
 XX 301 AGCCACTGAGGCTGCTCAGGCGACATCTGCCACAGCTGTGCTTACGCCCCGCTCC 360
 XX
 XX 361 TGTATCATGAGACCAAGCTCAATGCGAGGTTGAGAGAAACAATGATTGTGCGCATCAATC 420
 XX 361 TGTATCATGAGACCAAGCTCAATGCGAGGTTGAGAGAAACAATGATTGTGCGCATCAATC 420
 XX
 XX 421 TTCGAGCTATCAACAAGACCTGTGTTACCCCGCTGTGAGTGCCTTCAGGCAAGTGAATC 480
 XX 421 TTCGAGCTATCAACAAGACCTGTGTTACCCCGCTGTGAGTGCCTTCAGGCAAGTGAATC 480

DB 421 TTCGAGCTATCAACAAGACCTGTGTTACCCCGCTGTGAGTGCCTTCAGGCAAGTGAATC 480
 QY 481 CTCTATTCGAGCTCCCTCTAGACCCGAGCCCAATGCATGCGATCTGTG 540
 DB 481 CTCTATTCGAGCTCCCTCTAGACCCGAGCCCAATGCATGCGATCTGTG 540
 QY 541 TCGATTGACAGGATTTAAGTGTGCTGTGCGCAACCTTACCTGCGCTTT 600
 DB 541 TCGATTGACAGGATTTAAGTGTGCTGTGCGCAACCTTACCTGCGCTTT 600
 QY 601 GGACAAACACACTTGACATATGACTATACAAAGCCGACGCGAGGAAGC 660
 DB 601 GGACAAACACACTTGACATATGACTATACAAAGCCGACGCGAGGAAGC 660
 QY 661 TAATCTCTTGTGTCGCGGAGAGAGTTGTCAAGTTTGTATGATGAGATC 720
 DB 661 TAATCTCTTGTGTCGCGGAGAGAGTTGTCAAGTTTGTATGATGAGATC 720
 QY 721 CTGCGGAGATTTAGGCTTCAATCAATTGAGTACCTCTATCTTTCTTC 780
 DB 721 CTGCGGAGATTTAGGCTTCAATCAATTGAGTACCTCTATCTTTCTTC 780
 QY 781 AGCGAGTGCATCAAGAGCTGCGC 804
 DB 781 AGCGAGTGCATCAAGAGCTGCGC 804
 XX
 XX RESULT 3
 XX ABS70614
 XX ID ABS70614 standard; cDNA; 1531 BP.
 XX
 XX ABS70614;
 XX
 XX 28-NOV-2002 (first entry)
 XX
 XX Transcription factor stress-related protein (TFSRP), parti
 XX
 XX Transcription factor stress-related protein; TFSRP; MYB-1
 XX AP2 similar-2 protein; APS-2; zinc-finger factor-2; ZF-2;
 XX ZF-4 protein; ZF-5 protein; CAAT-box binding factor protei
 XX Sigma factor like protein; SFL-1; transgenic; plant; drouc
 XX environmental stress; Physcomitrella patens; high salinity
 XX ciliate; fungus; gene; ss.
 XX
 XX Physcomitrella patens.
 XX
 XX US2002102695-A1.
 XX
 XX 01-AUG-2002.
 XX
 XX 06-APR-2001; 2001US-00828303.
 XX
 XX 07-APR-2000; 2000US-0196001P.
 XX
 XX (SILV/) SILVA O D C E.
 XX (BOHN/) BOHNERT H J.
 XX (THIE/) THIELEN N V.
 XX (CHEN/) CHEN R.
 XX
 XX Silva ODCE, Bohnert HJ, Thielén NV, Chen R;
 XX MPI; 2002-690614/74.
 XX
 XX Novel transcription factor stress-related protein such as
 XX protein, zinc-finger protein, or CAAT-Box binding factor p
 XX useful for modifying stress tolerance of plant.
 XX
 XX Example 5; Fig 1D; 104pp; English.
 XX
 XX The invention relates to an isolated transcription factor
 XX protein (TFSRP) (i) selected from AP2 similar-2 protein (i
 XX finger factor-2 (ZF-2) protein, ZF-3 protein, ZF-4 protein;

MYB-1 protein, CAAT-Box binding factor protein-3 (CABF-3), Sigma factor like protein (SPF-1) and their orthologues. Also described is: (1) an isolated TFSRP coding nucleic acid (II) coding for (I); (2) a transgenic plant cell (III) transformed by (II), where the expression of (II) in the plant cell results in increased tolerance to an environmental stress as compared to a wild-type variety of the plant cell; (3) a transgenic plant (IV) comprising (III); (4) a seed (V) produced by (IV), where the seed is true breeding for an increased tolerance to environmental stress as compared to a wild-type variety of the plant cell; (5) an agricultural product (VI) produced by (IV) or (V); and (6) an isolated recombinant expression vector (VII) comprising (II), where expression of (VII) in a host cell results in increased tolerance to environmental stress as compared to a wild-type variety of the host cell. (II) is useful for modifying (increasing or decreasing) stress tolerance of a plant (transgenic or not transgenic) by modifying the expression of TFSRP in the plant. The plant is transformed with a promoter that directs expression of the TFSRP. The promoter is tissue specific and is developmentally regulated. TFSRP expression is modified by administration of an antisenese molecule that inhibits expression of TFSRP. (VII) is useful for producing a transgenic plant containing (II), where expression of the nucleic acid in the plant results in increased tolerance to environmental stress as compared to a wild-type variety of the plant, by transforming a plant cell with (VII) comprising the nucleic acid, generating from the plant cell a transgenic plant with an increased tolerance to environmental stress as compared to a wild-type variety of the plant. (I) is useful for conferring stress tolerance such as drought, cold and/or salt tolerance to plants. (I), (II), (III) or (VII) is useful for identifying phycomitrella patens and related organisms, for mapping of genomes of organisms related to P.patens, for identifying and localising P.patens sequences of interest, for evolutionary studies, for determining TFSRP regions required for function, for modulating TFSRP activity, for modulating metabolism of one or more cell functions, for modulating transmembrane transport of one or more compounds, and for modulating stress resistance. (I) is useful for reducing stress tolerance response plants or more particularly, in the transcription of a protein involved in a stress tolerance response in a P.patens plant. (II) is useful for transforming plants and thus inducing tolerance to stresses such as drought, high salinity and cold, for identifying the presence of P.patens or a related organism in a mixed population of microorganism, serve as marker for specific regions of the genome for mapping the genome and for the functional studies of P.patens proteins. (II) is useful as reference points for mapping the moss genomes, or of genomes of related organisms, for evolutionary and protein structural studies, for generating knockout mutation in the genomes of various organisms such as bacteria, mammalian cells, yeast cells and plant cells, useful for evaluating their ability or capacity to tolerate various stress conditions and the effect on the phenotype and/or genotype of the mutation. (II) is useful as marker for the construction of a genomic map in related mosses. (I) or (II) is useful for generating algae, ciliates, plants, fungi or other microorganisms expressing mutated TFSRP nucleic acid and protein molecules for improving stress tolerance. AB570611-AB570661 represent P. patens TFSRP coding sequences and PCR primers of the invention

Sequence 1531 BP, 313 A; 399 C; 393 G; 426 T; 0 U; 0 Other;

Query Match 96.6%; Score 776.4; DB 6; Length 1531;

Best Local Similarity 99.7%; Pred. No. 3.1e-241; Matches 788; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

8 GCACCAAGTCCCGCTTAGTGTGTGTCATTAAGTTGTTGCAAGCTGAAGCCTTGAGC 67
Db GCACCAAGTCCCGCTTAGTGTGTGTCATTAAGTTGTTGCAAGCTTGAGC 588
68 GAGATTGCGAGATTCTCATACGCTTGTGATTAGAAAGATACACCTTATTAGCTG 127
Db GAGATTGCGAGATTCTCATACGCTTGTGATTAGAAAGATACCTTATTAGCTG 648
128 TTAAGAGTGGCCACGAGGCTGTCTCAGAGAGACCTCGCAGGCCCTGAGGCTCA 187
Db TTAAGAGTGGCCACGAGGCTGTCTCAGAGAGACCTCGCAGGCCCTGAGGCTCA 708
188 GTTATGTGCAAGAACCTTTGCGGCTTCTTGCGACCAAGCTACATGAGGTTGTCTCG 247

|||||
Db GTTATGTGCAAGAACCTTTGCGGCTTCTTGCGACCAAGCTACATG 768
709 GTTATGTGCAAGAACCTTTGCGGCTTCTTGCGACCAAGCTACATG 307
248 AAGTGTCTCCAGAGACAGTATGCAAGCCGAAGTATGACGCTTTAGCTG 827
769 AAGTGTCTCCAGAGACAGTATGCAAGCCGAAGTATGACGCTTTAGCTG 367
308 CAGGCTGTCCAGAGACAGTATGCAAGCCGAAGTATGACGCTTTAGCTG 887
828 CAGGCTGTCCAGAGACAGTATGCAAGCCGAAGTATGACGCTTTAGCTG 427
368 GAGACCAAGCTCAGATGAGAGGTTGAGAGAACATGATGTGCGCATC 947
888 GAGACCAAGCTCAGATGAGAGGTTGAGAGAACATGATGTGCGCATC 487
428 TATCAACAAGCTGTGTTACCCCGCTGACAGTCCCTCAGGAGTA 1007
948 TATCAACAAGCTGTGTTACCCCGCTGACAGTCCCTCAGGAGTA 547
488 GCAGTCCCTCTAGACCCGAGCCCAATGATGCGGATTTGACAGAAC 1067
1008 GCAGTCCCTCTAGACCCGAGCCCAATGATGCGGATTTGACAGAAC 607
548 ACAGATTAAAGTGTGCTGTGCAACCTCTTACCTTACATGAGT 1127
1068 ACAGATTAAAGTGTGCTGTGCAACCTCTTACCTTACATGAGT 667
608 CACACTTCAATATGACTACAAAGCCGACAGGACGAGAGGATTTGCA 1187
1128 CACACTTCAATATGACTACAAAGCCGACAGGACGAGAGGATTTGCA 727
668 CTGTGTGCGGCGAGAGGTTGTCAAGTTTGTATGAGCATCCGTTAAC 1247
1188 CTGTGTGCGGCGAGAGGTTGTCAAGTTTGTATGAGCATCCGTTAAC 787
728 CGATTAGGCTTCAATTAAGTGAATCTTACATCTTTCTTTATAG 1307
1248 CGATTAGGCTTCAATTAAGTGAATCTTACATCTTTCTTTATAG 788
CGCATCAAGA 797
1308 CGCATCAAGA 1317

RESULT 4

AA595824 ID AA595824 standard; cDNA; 1531 BP.

AC AA595824;

DT 26-FEB-2002 (first entry)

DE Partial cDNA encoding Zinc finger-4 (ZF-4).

XX Transcription factor stress-related protein; AP2 similar-2

KM AP2-2; zinc-finger factor protein; ZF-2; ZF-3; ZF-4; ZF-5;

KM MYB-1 MYB-1 protein; CAAT-box binding factor protein-3; CAT

KM Sigma factor like protein; SPF-1; environmental stress; mo

OS Physcomitrella patens.

PN MO200177311-A2.

PD 18-OCT-2001.

PF 06-APR-2001; 2001MO-US011393.

PR 07-APR-2000; 2000US-0196001P.

PA (BADI) BASF PLANT SCI GMBH.

PI Da Costa Silva O, Bohnert HU, Van Thielén N, Chen R;

et;

XX DR MPI, 2002-049146/06.
XX PT New polypeptide, useful for identification of Physcomitrella patens, and
XX PT for modulating stress resistance of a plant, comprises an isolated
XX PT transcription factor stress-related protein derived from Physcomitrella
XX PT patens.
XX Example 5, Fig 1D, 129pp; English.
XX The invention relates to an isolated transcription factor stress-related
XX protein (TFSRP) (i) selected from an AP2 similar-2 protein (APS-2), a
XX zinc-finger factor-2 protein (ZF-2), ZF-3 protein, ZF-4 protein, ZF-5
XX protein, a MYB-1 protein (MYB-1), a CMT-box binding factor protein-3
XX (CABF-3), a Sigma factor like protein (SFL-1) and its orthologues. Also
XX described is an isolated TFSRP coding nucleic acid (ii) which codes for
XX (i); and an isolated recombinant expression vector (iii) comprising (ii),
XX where expression of (iii) in a host cell results in increased tolerance
XX to environmental stress as compared to a wild type variety of the host
XX cell. (i) or (ii) is useful as markers for specific regions of the TFSRP
XX genome. (i) or (ii) is useful for identification of Physcomitrella patens
XX and related organisms, mapping of genomes of organisms related to P.
XX patens, identification and localization of P. patens sequence of
XX interest, evolutionary studies, determination of TFSRP regions required
XX for function, modulation of a TFSRP activity, modulation of the
XX metabolism of one or more cell functions, modulation of the transmembrane
XX transport of one or more compounds and modulation of stress resistance.
XX (ii) is useful for generating probes and primers for identifying and/or
XX cloning TFSRP homologues in other cell types and other organisms, as well
XX as TFSRP homologues from mosses and related species, and for evolutionary
XX and protein structural studies. AAS95821-AAS95891 represent P. patens
XX TFSRP coding sequences and PCR primers of the invention
XX
SQ Sequence 1531 BP, 313 A; 399 C; 393 G; 426 T; 0 U; 0 Other;
Query Match 96.6%; Score 776.4; DB 6; Length 1531;
Best Local Similarity 99.7%; Pred No. 3.1e-241;
Matches 788; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 8 GCACGAGTCCGCTTAGTGTGTGTCATTAGTGTGTGCAAGTCTGAGCCTTGAGC 67
DB 529 GCACGAGTCCGCTTAGTGTGTGTCATTAGTGTGTGCAAGTCTGAGCCTTGAGC 588
QY 68 GAGATTGGAGAGATTTTCTCATACGCTTCTGATTAGGAAGATACACCTTATTAGTCTG 127
DB 589 GAGATTGGAGAGATTTTCTCATACGCTTCTGATTAGGAAGATACACCTTATTAGTCTG 648
QY 128 TTAAGATGGCCACGAGGCTGTGTCTCAGAGACGAGCTCGAGGCCCCCTGAGGGTCCA 187
DB 649 TTAAGATGGCCACGAGGCTGTGTCTCAGAGACGAGCTCGAGGCCCCCTGAGGGTCCA 708
QY 188 GTTATGTGCAAGAACCTTTTCGCGCTTCTTCGCGACGCACTTACATGAGGTTGCTCG 247
DB 709 GTTATGTGCAAGAACCTTTTCGCGCTTCTTCGCGACGCACTTACATGAGGTTGCTCG 768
QY 248 AAGTGTACCGAGAGACATCTATGCAAGCGGAAGATGACGGCTTTAGCTGAGCAAGCACT 307
DB 769 AAGTGTACCGAGAGACATCTATGCAAGCGGAAGATGACGGCTTTAGCTGAGCAAGCACT 827
QY 308 CAGGCTGTGCTCAGGCGCATCTGCAAGCTGTGCTGCTTCAAGCCCCCGCTCTGATCAT 367
DB 828 CAGGCTGTGCTCAGGCGCATCTGCAAGCTGTGCTGCTTCAAGCCCCCGCTCTGATCAT 887
QY 368 GAGACCAAGCTCACAATGCGAGGTTGAGAGAACATATGTCGCGATCAATCTTCCAGC 427
DB 888 GAGACCAAGCTCACAATGCGAGGTTGAGAGAACATATGTCGCGATCAATCTTCCAGC 947
QY 428 TATCAACAAGAGCTGTGTTACCCCGCTGAGCTGCGCCCTCAGGCAAGTGAATCTCTATC 487
DB 948 TATCAACAAGAGCTGTGTTACCCCGCTGAGCTGCGCCCTCAGGCAAGTGAATCTCTATC 1007
QY 488 GCAGCTCCCTTAGACCCGAGCCCAATGATGCGGATCTTGACGAGGAAGGTGTGGAATTG 547

DB 1008 GCAGCTCCCTTAGACCCGAGCCCAATGATGCGGATCTTGACGAGAGC 1067
QY 548 ACAGGATTTTAAGTGTGCTGTGTGGAACCTCTATCTGCGCTTTTACATCGGT 607
DB 1068 ACAGGATTTTAAGTGTGCTGTGTGGAACCTCTATCTGCGCTTTTACATCGGT 1127
QY 608 CACACTTGACATATGACTACGAAAGCCGAGGCGAGGAAGCGATTCGCA 667
DB 1128 CACACTTGACATATGACTACGAAAGCCGAGGCGAGGAAGCGATTCGCA 1187
QY 668 CTGTGTGTCGCGAGAGAGGTTGTCAGATTGATGAGCATCCGTTAAGC 727
DB 1188 CTGTGTGTCGCGAGAGAGGTTGTCAGATTGATGAGCATCCGTTAAGC 1247
QY 728 CGATTAGGCTTCATATCATATGAGTACTCTACATCTTCTCTTATCG 787
DB 1248 CGATTAGGCTTCATATCATATGAGTACTCTACATCTTCTCTTATCG 1307
QY 788 CGCATCAAGA 797
DB 1308 CGCATCAAGA 1317
RESULT 5
ADX47821
ID ADX47821 standard; cDNA; 1003 BP.
AC ADX47821;
XX 21-APR-2005 (first entry)
XX
XX DE Plant full length insert polynucleotide seqid 22561.
XX
XX KW plant protectant; plant growth regulant; gene therapy; plant
XX recombinant DNA construct; physical array; plant breeding;
XX cold tolerance; heat tolerance; drought tolerance; herbicide
XX extreme osmotic condition; pathogen tolerance; pest tolerance
XX growth rate; cell cycle pathway; disease resistance;
XX galactomannan production; lignin production; plant growth
XX yield; plant growth; plant development; seed oil; protein
XX protein content; gene; ss.
OS Unidentified.
XX
XX PN US2004034888-A1.
XX
XX PD 19-FEB-2004.
XX
XX PF 28-APR-2003; 2003US-00425114.
XX
XX PR 06-MAY-1999; 99US-00304517.
XX
XX PR 05-NOV-2001; 2001US-00985678.
XX
XX PA (LILUJ/) LILU J.
XX (ZHOUL/) ZHOU Y.
XX (KOVA/) KOVALIC D K.
XX (SCRE/) SCREEN S E.
XX (TABAR/) TABASKA J E.
XX (CAOY/) CAO Y.
XX
XX PI Liu J, Zhou Y, Kovacic DK, Screen SE, Tabaska JE, Cao
XX
XX DR MPI, 2004-180133/17.
XX
XX PT New recombinant DNA construct, useful for improving plant
XX cold, heat, drought, herbicide, extreme osmotic condition
XX pests, for conferring increased resistance to plant diseases
XX improving yield.
XX
XX PS Claim 1; SEQ ID NO 22561; 15pp; English.
XX
XX CC The invention describes a recombinant DNA construct comprising
XX polynucleotide consisting of a sequence encoding an amino

CC available in electronic form from the US patent office at
 CC ftp://seeddata.uspto.gov/sequence.html?DocID:2004034888. The polynucleotide
 CC of the invention are also useful in physical arrays of molecules and as
 CC plant breeding markers. The recombinant DNA construct is useful for
 CC improving plant tolerance to cold, heat, drought, herbicides, extreme
 CC osmotic conditions, pathogens or pests, for manipulating growth rate in
 CC plant cells by modification of the cell cycle pathway, for conferring
 CC increased resistance to plant disease, for producing galactomannan,
 CC lignin or plant growth regulators, for increasing the rate of homologous
 CC recombination in plants, for improving yield by modification of
 CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
 CC or by providing improved plant growth and development under at least one
 CC stress condition or for modifying seed oil or protein yield and/or
 CC content. This sequence represents a plant full length insert
 CC polynucleotide that can be used in the recombinant DNA construct of the
 CC invention.

XX Sequence 1003 BP; 211 A; 266 C; 294 G; 232 T; 0 U; 0 Other;

Query Match 12.9%; Score 103.4; DB 13; Length 1003;
 Best Local Similarity 53.6%; Pred. No. 2.1e-22;
 Matches 302; Conservative 0; Mismatches 231; Indels 30; Gaps 3;

QY 156 AGGAGACGACCTCGAGGCCCCCTGAGGCTCAGTATGTGCAAGAACCTTTGGGCTTCT 215
 DB 122 AGGAGGCTGGGTGCGAGGCCCCCGAGGAGCCATCTCTGTGATCAATAGTGGCTTCT 181
 QY 216 TCGGACGCGAAGCTACCATGAGGCTGTGCTGCAAGTCTGACGAGAGACATGCAAG 275
 DB 182 TCGGACGCGCGCGACCATGAAATGCTGCTCCAGTGGCCACAGAGATGATTAAGAAAGC 241
 QY 276 CGAAGATGACGCGCTTATGCTGAGCAAGCCACTCAGGCTGCTGAGGCGCATGTGCAAG 335
 DB 242 AGGATCAGGCGCAGCTGGCTGC-----CTCTTATTCAGACATGTGTA 287
 QY 336 CTGCTGCTGTTGAGCCCCCCCCCTCTCTGTAATGAGACCAAGCTCAGTGCAGGTTGAGA 395
 DB 288 CGGACGCGACGCGCT-----CATGAGCGCGTGTGCTGGAGCAAGACG 333
 QY 336 GAAACAATGATTTGCGCCGATCATCTTCAGCTATTCACAGAACCTGGTTACCCCGCTG 455
 DB 334 GATGATGCTGTTGCGCCAGTTGATGCAAAATGAAC--GTGACAGAGCCCGCTGATG 391
 QY 456 CAGCTGCCCTCAGGAGTGAAGTCTCTATGAGAGCTCCCTCTGACCCGAGCCCAATC 515
 DB 332 TTGCGGACCCAGGAGGCGGTGGCGGCAATTCAAAGGGGGAAGTAGGCGCAAC 451
 QY 516 GATGCGATCTTTGACAGAACCGTGTGATTTGACAGATTTAAAGTGTGCTGTGCAAC 575
 DB 452 GGTGACGCGCTCGAGGAAGAGGTTGACTTACGGAAATCACTGCCGGTGTGGAACT 511
 QY 576 TCTACTGCGCTTATCATGCTGTGAGCAAAACACATCTTGCAATGACTTGCAAGCCG 635
 DB 512 TGTACTGCGCACTTCACCGCTACTCCGACACAGACGATGCAAGTTGCACTCCGACTG 571
 QY 636 CAGGCGAGAGCGATTTGCGAAAGCTAATCTTGTGTGCGCGAGAGAGTTGTCAAGT 695
 DB 572 CTGCGAGGAGCGCAATTCGCAAGCTAATCCGCTGTGAGGACAGAGCTGCAAGAA 631
 QY 696 TTTGATGAGCATCCGTTAAGCTT 718
 DB 632 TCTAGGGGGGGTCTTACGGTT 654

RESULT 6
 ADX53121
 ID ADX53121 standard; cDNA; 873 BP.
 XX
 AC ADX53121;
 XX
 DT 21-APR-2005 (first entry)
 XX
 DE Plant; full length insert polynucleotide seqid 27861.

XX plant proectant; plant growth regulant; gene therapy; pla
 KW recombinant DNA construct; physical array; plant breeding
 KW cold tolerance; heat tolerance; drought tolerance; herbici
 KW extreme osmotic condition; pathogen tolerance; pest tolera
 KW growth rate; cell cycle pathway; disease resistance;
 KW galactomannan production; lignin production; plant growth
 KW yield; plant growth; plant development; seed oil; protein
 KW protein content; gene; ss.

XX Unidentified.

XX US2004034888-A1.

XX 19-FEB-2004.

XX 28-APR-2003; 2003US-00425114.

XX 06-MAY-1999; 99US-00304517.

XX 05-NOV-2001; 2001US-00985678.

XX (LIU/J) LIU J.

XX (ZHOU/J) ZHOU Y.

XX (KOVA/J) KOVALIC D K.

XX (SCRE/J) SCREEN S E.

XX (TABAJ) TABASKA J E.

XX (CAOY) CAO Y.

XX Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao

XX WPI; 2004-180133/17.

XX Claim 1; SEQ ID NO 27861; 15pp; English.

CC The invention describes a recombinant DNA construct compri
 CC polynucleotide consisting of a sequence encoding an amino
 CC available in electronic form from the US patent office at
 CC ftp://seeddata.uspto.gov/sequence.html?DocID:2004034888. The
 CC of the invention are also useful in physical arrays of mol
 CC plant breeding markers. The recombinant DNA construct is u
 CC improving plant tolerance to cold, heat, drought, herbicid
 CC osmotic conditions, pathogens or pests, for manipulating g
 CC plant cells by modification of the cell cycle pathway, for
 CC increased resistance to plant disease, for producing galact
 CC lignin or plant growth regulators, for increasing the rate
 CC recombination in plants, for improving yield by modificati
 CC photosynthesis or carbohydrate, nitrogen or phosphorus use
 CC or by providing improved plant growth and development unde
 CC stress condition or for modifying seed oil or protein yield
 CC content. This sequence represents a plant full length inse
 CC polynucleotide that can be used in the recombinant DNA con
 CC invention.

XX Sequence 873 BP; 180 A; 251 C; 264 G; 178 T; 0 U; 0 Other;

Query Match 12.8%; Score 103; DB 13; Length 87.
 Best Local Similarity 53.2%; Pred. No. 2.7e-22;
 Matches 307; Conservative 0; Mismatches 240; Indels

QY 156 AGGAGACGACCTCGAGGCCCCCTGAGGCTCAGTATGTGCAAGAACCT; 215
 DB 144 AGGAGGCTGGGTGCGAGGCCCCCGAGGAGCCATCTCTGTGATCAATTA 203
 QY 216 TCGGACGCGAAGCTACCATGAGGCTGTGCTGCAAGTGTACCGAGAGAC 275
 DB 204 TCGGACGCGCGCGACCATGAAATGCTGCTTATGCTGCAAGAGAT; 263
 QY 276 CGAAGATGACGCGCTTATGCTGAGCAAGCCACTCAGGCTGCTCAGGCGAC 335

DB 264 AGGATCAGGCCAGCGCTGC-----CTCTCTATGACAGCATCTGTA 309
QY 336 CTGCTGCTGTTGAGCCCCCGCTCTGTATCATGAGCAAGCTCATGCGAGTTGAGA 395
DB 310 CGGACGCGACGCCGT-----CATGAGCCGGTTGTTGCTGGACAAACAG 355
QY 396 GAAACATGATTTGTGCGCGCATCATCTTCCAGCTATTAACAAGACCTGGTTACCCCGCTG 455
DB 356 GATGATAGCTGTGGCCCAAGTCAGATTGCAAAACATGAMC--GTGACGACGCCCGCTGATG 413
QY 456 CAGCTGCCCTCAGGACGATGAAGTCTCTATGCGACGCTCCCTCTAGAACCCGAGCCCAATC 515
DB 414 TTGCCGACCCAGCCAGAGGGGGTGGCGGCATCTCCAAAGGGGGAAGTAGGGCCGACCC 473
QY 516 GATGCGGATCTTTCAGAGAAAGCGTTGTGATTAAGATTTAAGTGTGCTGGCAACC 575
DB 474 GGTGACGCGCTGACAGAAAGAGGGTGGGCTCACGGGATTTCAACTGCCGTGGGAACT 533
QY 576 TCTACTGCGCTTACATCGGTACTCTGGAAACACACTTGCACATATGACTACAAAGCCG 635
DB 534 TGTACTGCGCGCTCCACCGCTACTCGACACAGACGACTGCAAGTTGACTCCGACCTG 593
QY 636 CAGGCGAGGAAGCGATTGCGAAAGCTAATCTTGTGCGTGGCCGAGAGTTGTCAAGT 695
DB 594 CTGCCAGGAGCGCCATTGCCAAAGCTAATCCGCTGTGAGGAGCAACAGCTCGACACAGA 653
QY 696 TTTGATGAGCATCCGTTAAGCTTTCTGCGCAGCAT 732
DB 654 TCTAGGGGGGGTCTCTAGCGTTGGTGTACAGAAAGT 690

RESULT 7

ADXS1406 standard; cDNA; 945 BP.

ADXS1406;

21-APR-2005 (first entry)

Plant full length insert polynucleotide seqid 26146.

plant protectant; plant growth regulant; gene therapy; plant;
recombinant DNA construct; physical array; plant breeding marker;
cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
extreme osmotic condition; pathogen tolerance; pest tolerance;
growth rate; cell cycle pathway; disease resistance;
galactomannan production; lignin production; plant growth regulator;
yield; plant growth; plant development; seed oil; protein yield;
protein content; gene; ss.

Unidentified.

US2004034888-A1.

19-FEB-2004.

28-APR-2003; 2003US-00425114.

06-MAY-1999; 99US-00304517.

05-NOV-2001; 2001US-00985678.

(LIU/J) LIU J.
(ZHOU/J) ZHOU Y.
(KOVA/J) KOVALIC D K.
(SCRE/J) SCREEN S E.
(TABAS/J) TABASKA J E.
(CAO/Y) CAO Y.

Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y,
WPI, 2004-180133/17.

PT New recombinant DNA construct, useful for improving plant
PT cold, heat, drought, herbicides, extreme osmotic condition
PT pests, for conferring increased resistance to plant disease
PT improving yield.

Claim 1, SEQ ID NO 26146; 15bp; English.

CC The invention describes a recombinant DNA construct compri
CC polynucleotide consisting of a sequence encoding an amino
CC available in electronic form from the US patent office at
CC ftp.segdata.ubpo.gov/sequence.html?docid:2004034888. The
CC of the invention are also useful in physical arrays of mol
CC plant breeding markers. The recombinant DNA construct is
CC improving plant tolerance to cold, heat, drought, herbicide
CC osmotic conditions, pathogens or pests, for manipulating
CC plant cells by modification of the cell cycle pathway, fo
CC increased resistance to plant disease, for producing galac
CC lignin or plant growth regulators, for increasing the rate
CC recombination in plants, for improving yield by modificati
CC photosynthesis or carbohydrate, nitrogen or phosphorus use
CC or by providing improved plant growth and development unde
CC stress condition or for modifying seed oil or protein yiel
CC content. This sequence represents a plant full length inse
CC polynucleotide that can be used in the recombinant DNA con
CC invention.

SQ Sequence 945 BP; 207 A; 245 C; 256 G; 237 T; 0 U; 0 Other;

Query Match

Best Local Similarity 12.8%; Score 103; DB 13; Length 94

Matches 284; Conservative 0; Mismatches 220; Indels

QY 156 AGGACGACACCTCGAGGCCCTGAGGGTTCAGATTATGTGCAAGACCT 215
DB 113 AGGACATGATGATGCGACACCTGAGGAGACCATCTTTGATCAATAT. 172
QY 216 TCGGACGCAAGCTACATGAGGGTGTGCTGGAATGTCTACCGAAGAC 275
DB 173 TCGGACGCGAGCTACATGAACATGTGCTCCAGTGTGCACAAAGAGAT 232
QY 276 CGAAGATGACGGCTTTAAGCTGAGCAAGCACTCAGGTGTGACGCGAC 335
DB 223 AGGACGAGGCCAAGTGTGCTG----- 253
QY 336 CTGCTGCTGTTACGCCCGCTCTGTATCATGAGCAAGCTCATATG 395
DB 254 CTCTCTATGACACAGATCGTCATATGGAAGAGATGTGTATGGAAC 313
QY 396 GAAACATGATTTGTGCGCGCATATTTCCAGCTATCAACAAAGCTGTG 455
DB 314 GCACACAGTGTGCTCTGCTGCTCAATCGAGTTGCAACATGAACGT 373
QY 456 CAGTGGCCCTCAGGACGATGAAGTCTCTATGCGACGCTCCCTAGAC 515
DB 374 ATGTGCTGAGACTTAGCGAGAGGGGGGCGGTGATCTCCAAAGGAAAG 433
QY 516 GATGCGGATCTTTCAGAGAAAGCGTGTGATTTGACAGAAATTAAGTGT 575
DB 434 GGTGACGACCTTTCAGAGAAAGGGTGTGATTTACAGATTCAGATCGCC 493
QY 576 TCTACTGCGCTTTCATGCTGATCTGGAACAAACACTTGCACATATGA 635
DB 494 TGTACTGTGCACTGACCGCTACTCGCAAGACGACGACTGCAAGTTGCA 553
QY 636 CAGGCGAGGAAGCGATTGCGAAAGCTAATCTTGTGTGTCGCGAGAA 695
DB 554 CTGCTAGGGATGCCATTGCGAAGGCTAATCCAGTGTGAAGCGGACAA 613
QY 696 TTT 698
DB 614 TCT 616

RESULT 8
 ADX51034
 ID ADX51034 standard; cDNA, 932 BP.
 XX
 AC
 XX ADX51034;
 DT 21-APR-2005 (first entry)
 XX
 DE Plant full length insert polynucleotide seqid 25774.
 XX
 KW plant protectant; plant growth regulant; gene therapy; plant;
 KW recombinant DNA construct; physical array; plant breeding marker;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
 KW extreme osmotic condition; pathogen tolerance; pest tolerance;
 KW growth rate; cell cycle pathway; disease resistance;
 KW galactomannan production; lignin production; plant growth regulator;
 KW yield; plant growth; plant development; seed oil; protein yield;
 KW protein content; gene; ss.
 XX
 OS unidentified.
 XX
 PN US2004034888-A1.
 XX
 PD 19-FEB-2004.
 XX
 PF 28-APR-2003; 2003US-00425114.
 XX
 PR 06-MAY-1999; 99US-00304517.
 XX
 PR 05-NOV-2001; 2001US-00985678.
 XX
 PA (LIU/J) LIU J.
 PA (ZHOU/) ZHOU Y.
 PA (KOVA/) KOVALIC D K.
 PA (SCRE/) SCREEN S E.
 PA (TABA/) TABASKA J E.
 PA (CAO/Y) CAO Y.
 XX
 PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
 DR WPI; 2004-180133/17.
 XX
 PT New recombinant DNA construct, useful for improving plant tolerance to
 PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
 PT pests, for conferring increased resistance to plant disease, or for
 PT improving yield.
 XX
 PS Claim 1; SEQ ID NO 25774; 15pp; English.
 XX
 XX
 XX The invention describes a recombinant DNA construct comprising a
 CC polynucleotide consisting of a sequence encoding an amino acid sequence
 CC available in electronic form from the US patent office at
 CC ftp://seqdata.uspto.gov/sequence.html?docID:2004034888. The polynucleotide
 CC of the invention are also useful in physical arrays of molecules and as
 CC plant breeding markers. The recombinant DNA construct is useful for
 CC improving plant tolerance to cold, heat, drought, herbicides, extreme
 CC osmotic conditions, pathogens or pests, for manipulating growth rate in
 CC plant cells by modification of the cell cycle pathway, for conferring
 CC increased resistance to plant disease, for producing galactomannan,
 CC lignin or plant growth regulators, for increasing the rate of homologous
 CC recombination in plants, for improving yield by modification of
 CC photoassimilates or carbohydrate, nitrogen or phosphorus use and/or uptake
 CC or by providing improved plant growth and development under at least one
 CC stress condition or for modifying seed oil or protein yield and/or
 CC content. This sequence represents a plant full length insert
 CC polynucleotide that can be used in the recombinant DNA construct of the
 CC invention.
 XX
 XX Sequence 932 BP; 205 A; 236 C; 264 G; 227 T; 0 U; 0 Other;

QY	82	TTTCTCATACGCTTCTGTATAGAAAGATACACCCCTTATTAAGTCTGTAA	1414
Db	136	TGTCACTACTATATCTGGATTGAATGACAGTGTGTTTGGTATTAATCTAGATT	1959
QY	142	CGAGCGTGTGTCTCAGAGACGACCTTCGCAAGGCCCTCGAAGGATCCAGTTT	2010
Db	196	GGAGC-----ACAAAGAGGCGTGGGTGCCAAGGCCCCCGAAGGGACCATTC	2494
QY	202	CCTTTGGCGCTTCTTCGGCAGCCAAAGTAACTCATGGGGTTGTGTCCAGAG	2614
Db	250	TAACTGTGGCTTCTTCGGCAGGCGCGCGACATGAACAATGTGCTCCAGG	3090
QY	262	GACAGTCATGCAAGCGAAGATGACGGCTTTAACTGAGCAAGCAACTCAG	3212
Db	310	GATGATTAACGAAGCAGAGATCAGGCCAAGCTGGCTGC-----	3555
QY	322	GACATCTGCCACACAGCTGCTGTCTGTTCAAGCCGCCGCTCCTGTATATAG	3811
Db	356	GACACGATCGTGAAACGGCAGCGACGCCGT-----CATGGA	4010
QY	382	ATGCGAGGTTGAGAGAAACAATGATTTGTGCCGATCAATCTTCAGGTTAT	4411
Db	402	CTGGCAGCAACACGGTAGTAGCTGTGGCCCAAGTTGAGTTGCACAAACAT	4559
QY	442	GATTACCCCGCTGCAGCTGCCCCCTCAGGCAAGTGAAGTCCCTATGCGCA	5010
Db	460	GCAGCCGCTGATGTGTTGCCGACCCAGAGGAGGGGGTGGCGGAGATCTCC	5119
QY	502	ACCCGAGCCCAATCGATGCGGATCTTTCGACGAAAGCGTGTGGATTACACA	5611
Db	520	GGTAAAGGCCGAACCGGTGCAGCGCTTCGACGAAAGAGGGTTGACCTTAACG	5797
QY	562	TCGCTGTGGCAACCTCTACTGCGCTTTTACATCGGTACTCGACAAACAC	6212
Db	580	CCGGTGTGGAACTTGTACTGGCACTCCACCGCTACTCCGACAAACAC	6399
QY	622	TGACTACAAAGCCGACGGCAGAGAACGATTCGAAAGCTTAATCCTCTT	6811
Db	640	CGACTACCGGACTGCTGCAGGGAAGCCCATTCCAAGGCTTAATCCCGTGC	6999
QY	682	GAAGGTGTCAAGTTTGTATGAGCANCGTTTAAAGCTT 718	
Db	700	CAAGCTCGACAAAGATCTAAGGGGGGTTCCCTTACGGTT 736	
RESULT 9			
ADX52834			
ID	ADX52834	standard; cDNA; 1032 BP.	
XX	ADX52834;		
AC			
XX			
DT	21-APR-2005	(first entry)	
XX			
DE	Plant full length insert polynucleotide seqid 27574.		
XX			
KM	plant protectant; plant growth regulant; gene therapy; pla		
KM	recombinant DNA construct; physical array; plant breeding		
KM	cold tolerance; heat tolerance; drought tolerance; herbici		
KM	extreme osmotic condition; pathogen tolerance; pest tolera		
KM	growth rate; cell cycle pathway; disease resistance;		
KM	galactomanan production; lignin production; plant growth		
KM	yield; plant growth; plant development; seed oil; protein		
KM	protein content; gene; ss.		
OS	unidentified.		
XX			
PN	US2004034888-A1.		
XX			
PD	19-FEB-2004.		
XX			
PF	28-APR-2003; 2003US-00425114.		
XX			
PR	06-MAY-1999; 99US-00304517.		

PR 05-NOV-2001; 2001US-00985678.
 XX (LIU/J) LIU J.
 PA (ZHOU/) ZHOU Y.
 PA (KOVA/) KOVALIC D K.
 PA (SCRE/) SCREEN S E.
 PA (TABAS/) TABASKA J E.
 PA (CAOY/) CAO Y.
 XX
 PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
 XX WPI, 2004-180133/17.
 XX
 PT New recombinant DNA construct, useful for improving plant tolerance to
 PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
 PT pests, for conferring increased resistance to plant disease, or for
 PT improving yield.
 XX
 PS Claim 1; SEQ ID NO 27574; 15pp; English.

CC The invention describes a recombinant DNA construct comprising a
 CC polynucleotide consisting of a sequence encoding an amino acid sequence
 CC available in electronic form from the US patent office at
 CC ftp.segdata.uspto.gov/sequence.html?docid:2004034888. The polynucleotide
 CC of the invention are also useful in physical arrays of molecules and as
 CC plant breeding markers. The recombinant DNA construct is useful for
 CC improving plant tolerance to cold, heat, drought, herbicides, extreme
 CC osmotic conditions, pathogens or pests, for manipulating growth rate in
 CC plant cells by modification of the cell cycle pathway, for conferring
 CC increased resistance to plant disease, for producing galactomannan,
 CC lignin or plant growth regulators, for increasing the rate of homologous
 CC recombination in plants, for improving yield by modification of
 CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
 CC or by providing improved plant growth and development under at least one
 CC stress condition or for modifying seed oil or protein yield and/or
 CC content. This sequence represents a plant full length insert
 CC polynucleotide that can be used in the recombinant DNA construct of the
 CC invention.

XX Sequence 1032 BP; 211 A; 278 C; 304 G; 239 T; 0 U; 0 Other;

Query Match 12.7%; Score 102.2; DB 13; Length 1032;
 Best Local Similarity 51.2%; Pred. No. 5.3e-22;
 Matches 288; Conservative 0; Mismatches 248; Indels 27; Gaps 1;

QY 156 AGGAGACGACCTCGAGGCCCCCTGAGGCTCAGTATGTGCAAGAACTTTGGGCTTCT 215
 DB 135 AGGAGGCTGGGTGCGCAGGCCCCCGAGGGACCCTCTGTCATCAATACGTGGCTTCT 194
 QY 216 TCGGACGCGCAAGCTACATGAGGTTGTGTCGAAGTGTCAAGAGACATCATGCAAG 275
 DB 195 TCGGACGCGCGGCGACATGAACTATGTGCTTAACTGTGCACAAAGAGATGATMAAGAAC 254
 QY 276 CGAAGATGACGGCTTTAGCTGAGCAAGCCACTCAGAGCTGCTGAGCGCATCTGCACAG 335
 DB 255 AGGA-----TCAGGCCAAGCTGGCTGCTCTCTTATTTG 287
 QY 336 CTGCTGCTGTACGCCCCCGCTCTGTATCATGAGAACCAAGCTCATGCGAGTTGAGA 395
 DB 288 ACAGCATGCTGAACGCGACGCGCGCTCATGAGACCGGTTGCTGTGACACAAACAGG 347
 QY 396 GAACATATATTTGCGCCGATCAATCTTCAGCTATCAACAGAACTGGTTACCCCGCTG 455
 DB 348 TAGTAGCTGTGTGCCAAGTCGATGTTGCAAAACATGAACGTGACAGCCCGCTGATGTTG 407
 QY 456 CAGTGGCCCTCAGGACAGTAAGTCTCTATCGAGCTCCCTCTAGACCCGAGCCCAATC 515
 DB 408 CCGAACCCACGACGAGGGGGTGGGGCGGCGATCTCCAAAGGGGGGAAGTAGGGCGAAAC 467
 QY 516 GATGCGGATCTTGGACAGAAAGCTGTTGATTGACAGATTAAAGTGTGCTGTGGCAAC 575
 DB 468 GGTGACGCGCTCGACGAAAGAGGTTGGGCTCACGGGATTCAAATGCGCGGTGTGGAACT 527

QY 576 TCTACTGGGCTTTACATGGTACTCGGACAAACACACTTGCATATAG.
 DB 528 TGTACTGGCGGCTTCCACGCTACTCGAACAACAGACTGCAAGTTGCA
 QY 636 CAGGCGAGGAAGCGATTGCGAAAGCTAACTCTTGTGTCGCCGAGAA.
 DB 588 CTGCGACGGAAGCGCATTTGCGAAAGCTAACTCGGCTGTGAAGCACAACA.
 QY 696 TTTGATGAGCATCCGTTAAGCTT 718
 DB 648 TCTAGGCGGCGTTCCCTACGGTT 670

RESULT 10
 ADX47307
 ID ADX47307 standard; cDNA; 736 BP.
 XX
 AC ADX47307;
 XX
 DT 21-APR-2005 (first entry)

XX Plant full length insert polynucleotide seqid 22047.
 XX
 DB plant protectant; plant growth regulant; gene therapy; plant
 XX recombinant DNA construct; physical array; plant breeding
 XX cold tolerance; heat tolerance; drought tolerance; herbic;
 KW extreme osmotic condition; pathogen tolerance; pest toler;
 KW growth rate; cell cycle pathway; disease resistance;
 KW galactomannan production; lignin production; plant growth
 KW yield; plant growth; plant development; seed oil; protein
 KW protein content; gene; ss.

XX Unidentified.
 XX
 XX US2004034888-A1.
 XX
 XX 19-FEB-2004.

XX 28-APR-2003; 2003US-00425114.

XX 06-MAY-1999; 99US-00304517.
 PR 05-NOV-2001; 2001US-00985678.

PA (LIU/J) LIU J.
 PA (ZHOU/) ZHOU Y.
 PA (KOVA/) KOVALIC D K.
 PA (SCRE/) SCREEN S E.
 PA (TABAS/) TABASKA J E.
 PA (CAOY/) CAO Y.

PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao
 DR WPI, 2004-180133/17.

PT New recombinant DNA construct, useful for improving plant
 PT cold, heat, drought, herbicides, extreme osmotic condition
 PT pests, for conferring increased resistance to plant disease
 PT improving yield.

PS Claim 1; SEQ ID NO 22047; 15pp; English.

CC The invention describes a recombinant DNA construct compri
 CC polynucleotide consisting of a sequence encoding an amino
 CC available in electronic form from the US patent office at
 CC ftp.segdata.uspto.gov/sequence.html?docid:2004034888. The
 CC of the invention are also useful in physical arrays of mol
 CC plant breeding markers. The recombinant DNA construct is u
 CC improving plant tolerance to cold, heat, drought, herbicid
 CC osmotic conditions, pathogens or pests, for manipulating g
 CC plant cells by modification of the cell cycle pathway, for
 CC increased resistance to plant disease, for producing galac
 CC lignin or plant growth regulators, for increasing the rate
 CC recombination in plants, for improving yield by modificati

CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC or by providing improved plant growth and development under at least one
CC stress condition or for modifying seed oil or protein yield and/or
CC content. This sequence represents a plant full length insert
CC polynucleotide that can be used in the recombinant DNA construct of the
CC invention.

XX Sequence 736 BP; 158 A; 213 C; 225 G; 140 T; 0 U; 0 Other;

XX Query Match 12.7%; Score 101.8; DB 13; Length 736;

XX Best Local Similarity 53.5%; Pred. No. 6e-22; Matches 301; Conservative 0; Mismatches 232; Indels 30; Gaps 3;

QY 156 AGGAGACGACCTCGAGGCCCCCTGAGGGTCCAGTTATGCGAAGAACCTTTGCGGCTTCT 215
DB 114 AGGAGGCTGGGTCCAGGCCCCCGAGAGGCCATCTCTGCAATCAATTAATCTGGCTTCT 173
QY 216 TCGGAGCCCAAGCTACATGGGGTTGTGCTCGAAGTGTCTACCGAGAGACATGCAAG 275
DB 174 TCGGAGGCGCGCGACATGAACATGTGCTCCAGTGGCCACAGAGATGATTAAGAAAC 233
QY 276 CGAAGATGACGGCTTTAGCTGAGCAAGCCACTCAGGCTGCTCAGCGCATCTGCCAAG 335
DB 234 AGGATCAGGCCCAAGCTGGCTGC-----CTCCTCTATCGACAGCATGTGA 279
QY 336 CTGCTGCTGTTGAGCCCCCGCTCCGTATCATGAGACCAAGCTCAATGCGAGTTGAGA 395
DB 280 CGGACGCCAGCCGT-----CATGAGCCCGGTTGTTGCTGGCAGCAACAG 325
QY 396 GAACATGATTTGTGCGCATCATCTTCCAGCTATCAACMAAGCCTGTTACCCCGCTG 455
DB 336 GTAGTAGCTGTGGCCCAAGTTGCAAAACATGAAC--GTGACAGACGCCCGCTGATG 383
QY 456 CAGCTGCCCTTAGGCAATGAAGTCTCTATGCAAGCTCCCTCTGACCCCGCAATC 515
DB 384 TTGCCCGAACCCAGCGAGGGGTGGCGGCATCTCCMAAGGGGGAAGTAGGGCCGAAC 443
QY 516 GATGCGGATCTTGGACGAAGCGGTGTGATGACAGGATTTAAGTGTGCTGGCAAC 575
DB 444 GGTGACGCCCTCGACGAAGAGGGTTGACTTACGGGATTTCACTGCCGTGGGAACT 503
QY 576 TCTACTGCGCTTTAGATCGGTACTCGGACMAACACTTGCACATATGACTACMAAGCG 635
DB 504 TGTACTGGGCACTCCACCGCTACTCCGACMAACAGACTGCAAGTTGCACTACGGAGCTG 563
QY 636 CAGGCGAGGAAGCGATTTGGGAAAGCTAATCTCTTGTCTGTGCGGACGAAGGTTTCAAGT 695
DB 564 CTGCCAGGGAGCCCATTTGCCAAGGCTAATCCGCTGGTGAAGGACAGCAAGCTCGACAAGA 623
QY 696 TTTGATGAGCATCCGTTAAGCTT 718
DB 624 TCTAGGGGGGTTCCCTACGCTT 646

RESULT 11

ADXS3784 ID ADXS3784 standard; cDNA; 776 BP.

XX ADXS3784;

XX 21-APR-2005 (first entry)

XX Plant full length insert polynucleotide seqid 28524.

XX plant protectant; plant growth regulant; gene therapy; plant;
XX recombinant DNA construct; physical array; plant breeding marker;
XX cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
XX extreme osmotic condition; pathogen tolerance; pest tolerance;
XX growth rate; cell cycle pathway; disease resistance;
XX galactomannan production; lignin production; plant growth regulator;
XX yield; plant growth; plant development; seed oil; protein yield;
XX protein content; gene; ss.

OS unidentified.

XX US2004034888-A1.

XX 19-FEB-2004.

XX 28-APR-2003; 2003US-00425114.

XX 06-MAY-1999; 99US-00304517.

XX 05-NOV-2001; 2001US-00985678.

XX (LIU/J.) LIU J.

XX (ZHOU/Y.) ZHOU Y.

XX (KOVA/) KOVALIC D K.

XX (SCRE/) SCREEN S E.

XX (TABBA/) TABASKA J E.

XX (CAO/Y.) CAO Y.

XX Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao

XX WPI, 2004-180133/17.

XX Claim 1; SEQ ID NO 28524; 15bp; English.

CC The invention describes a recombinant DNA construct compris
CC polynucleotide consisting of a sequence encoding an amino
CC available in electronic form from the US patent office at
CC ftp://seeddata.uspto.gov/sequence.html?docid:2004034888. The i
CC of the invention are also useful in physical arrays of mol
CC plant breeding markers. The recombinant DNA construct is u
CC improving plant tolerance to cold, heat, drought, herbicid
CC osmotic conditions, pathogens or pests, for manipulating g
CC plant cells by modification of the cell cycle pathway, for
CC increased resistance to plant disease, for producing galac
CC lignin or plant growth regulators, for increasing the rate
CC recombination in plants, for improving yield by modificati
CC photosynthesis or carbohydrate, nitrogen or phosphorus use
CC or by providing improved plant growth and development unde
CC stress condition or for modifying seed oil or protein yield
CC content. This sequence represents a plant full length inse
CC polynucleotide that can be used in the recombinant DNA con
CC invention.

XX Sequence 776 BP; 162 A; 225 C; 238 G; 151 T; 0 U; 0 Other;

XX Query Match 12.7%; Score 101.8; DB 13; Length

XX Best Local Similarity 53.5%; Pred. No. 6.2e-22;

XX Matches 301; Conservative 0; Mismatches 232; Indels

QY 156 AGGAGACGACCTCGAGGCCCCCTGAGGGTCCAGTTATGCGAAGAACCT 215
DB 134 AGGAGGCTGGGTCCAGGCCCCCGAGAGGCCATCTCTGCAATCAATTA 193
QY 216 TCGGAGCCCAAGCTACATGGGGTTGTGCTCGAAGTGTCTACCGAGAGC 275
DB 194 TCGGAGGCGCGCGACCACTGAACATGTGCTTAAGTGGCCACAAGAGAT 253
QY 276 CGAAGATGACGGCTTTAGCTGAGCAAGCACTCAGGCTGCTCAGGCGAG 335
DB 254 AGGATCAGGCCCAAGCTGGCTGC-----CTCCTCTATTCGAC 299
QY 336 CTGCTGCTGTTCAAGCCCCCGCTCCTGTACATGAGACCAAGCTCAGAG 395
DB 300 CGGACGCCAGCCGT-----CATGAGCCCGGTTGTTGCTG 345
QY 396 GAACATGATTTGTGCGCATCAATCTTCCAGCTATCAACAAAGACTGCT 455
DB 346 GTAGTAGCTGTGGCCCAAGTGAAGTTGCAAAACATGAAC--GTGACGCA 403

QY 456 CAGCTGCCCTCAAGCAAGTAAGTCTCTATTCGAGCTCCCTCTAGACCCGAGCCCAATC 515
DB 404 TTGCCCCGACCCAGCGAGGGGGTGGCCGCGATCTCCAAAGGGGGAAGTGAAGCCGAAACC 463
QY 516 GATCGGATCTTTCAGGAAGCGTGTGATGACAGGATTTAAGTGTGCTGTGGCAACC 575
DB 464 GGTGCAAGCCCTTCAGGAAGAGGTTGGGCTCAGCGGATTCACCTGCGGTGGGAACT 523
QY 576 TCTACTGCGCTTTCATCGGTAATCGGACAAACACTTGCACATATGACTACAAAGCCG 635
DB 524 TGTACTGCGGCTCCACCGCTACTCCGACAAAGCAAGCAAGTTCGACTACCGGAACT 583
QY 636 CAGGCAAGAAAGGATTTGCAAAAGCTATCTCTGTGCTGCGCGGAGAAAGTTGCAAGT 695
DB 584 CTGCGCAGGAGCGCATTTGCCAAGCTAATCCGGTGTGAAGGACAAAGCTCAGCAAGA 643
QY 696 TTTGATGAGCATCCGTTAAGCTT 718
DB 644 TCTAGGGGGGTTCCCTACGGTT 666

RESULT 12

ADK47342

ID ADK47342 strand: cDNA; 821 BP.

XX ADK47342;

XX 21-APR-2005 (first entry)

XX Plant full length insert polynucleotide seqid 22082.

XX plant protectant; plant growth regulant; gene therapy; plant;
XX recombinant DNA construct; physical array; plant breeding marker;
XX cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
XX extreme osmotic condition; pathogen tolerance; pest tolerance;
XX growth rate; cell cycle pathway; disease resistance;
XX galactomannan production; lignin production; plant growth regulator;
XX yield; plant growth; plant development; seed oil; protein yield;
XX protein content; gene; ss.

XX Unidentified.

XX US2004034888-A1.

XX 19-FEB-2004.

XX 28-APR-2003; 2003US-00425114.

XX 06-MAY-1999; 99US-00304517.

XX 05-NOV-2001; 2001US-00985678.

XX (LIU/J) LIU J.

XX (ZHOU/) ZHOU Y.

XX (KOVA/) KOVALIC D K.

XX (SCRE/) SCREBEN S E.

XX (TABA/) TABASKA J E.

XX (CAOY/) CAO Y.

XX Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;

XX WPI; 2004-180133/17.

XX New recombinant DNA construct, useful for improving plant tolerance to
XX cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
XX pests, for conferring increased resistance to plant disease, or for
XX improving yield.

XX Claim 1; SEQ ID NO 22082; 15pp; English.

XX The invention describes a recombinant DNA construct comprising a
XX CC polynucleotide consisting of a sequence encoding an amino acid sequence
XX available in electronic form from the US patent office at

CC ftp.segdata.uspto.gov/sequence.html?DocID:2004034888. The
CC of the invention are also useful in physical arrays of mol
CC plant breeding markers. The recombinant DNA construct is u
CC improving plant tolerance to cold, heat, drought, herbicid
CC osmotic conditions, pathogens or pests, for manipulating g
CC plant cells by modification of the cell cycle pathway, for
CC increased resistance to plant disease, for producing galac
CC lignin or plant growth regulators, for increasing the rate
CC recombination in plants, for improving yield by modificati
CC photosynthesis or carbohydrate, nitrogen or phosphorus use
CC or by providing improved plant growth and development unde
CC stress condition or for modifying seed oil or protein yield
CC content. This sequence represents a plant full length ins
CC polynucleotide that can be used in the recombinant DNA con
CC invention.

SQ Sequence 821 BP; 171 A; 236 C; 247 G; 167 T; 0 U; 0 Other;

Query Match

Best Local Similarity 12.7%; Score 101.8; DB 13; Length
Matches 301; Conservative 0; Mismatches 232; Indels 3;

QY 156 AGGAGACGACCTCGCAGGCGCCCTGAGGCTCCAGTTATGCAAGAACT
DB 114 AGGAGGCTGGGTGCCAGGCTCCCGAGGAGCCATCTCTGCACTAATA
QY 216 TCGGACGCCAAGCTACATGAGGTTGTGCTCGAAGTCTACCGAGAGAC
DB 174 TCGGACGCGCGGAGACATGAACATGTGCTCCAAAGTCCCAAGAGAT
QY 276 CGAAGTACGCGCTTAACTGACGAAGCACTCAGGCTGCTCAGGAGAC
DB 234 AGGATCAGGCCAAGCTGCGTGC-----CTCCTCTATGAC
QY 336 CTGCTGCTGTTCAGGCGCCCGCTCCTGTCATGAGACCAAGCTCATG
DB 280 CCGACGACGACCGCT-----CATGAGCCGCGTTGTGCTG
QY 396 GAACAATGATGTGCGCATCAATCTTCAGCTATCAACAAGACTGTG
DB 326 GTATGATGCTGTGCCAAGTTGAGTTGCAACAATGAAAC--GTGACGAC
QY 456 CAGCTGCCCTCAGGACGTAAGTCTCTATGCGACGCTCCTCTAGAC
DB 384 TTGCGGAGCCAGCGAGGGGGTGGCGGCGATCTCCAAAGGGGGAAGGT
QY 516 GATGGGATCTTGAAGAGCGTGTGATTTGACAGGATTTAAGTGTG
DB 444 GGTGACGCGCTGCAAGAAAGGTTTGAATTACGGAATTCATGCGC
QY 576 TCTACTGCGCTTTCATCGGTAATCGGACAAACACTTGCACATATG
DB 504 TGTACTGCGCACTCCACGCTACTCCGACAAAGCAAGCTGCAAGTTG
QY 636 CAGGCAAGAAAGGATTTGCAAAAGCTAATCTCTGTGCTGCGCGAAG
DB 564 CTGCGCAGGAGCGCATTTGCCAAGCTAATCCGGTGTGAAGGACAGCA
QY 696 TTTGATGAGCATCCGTTAAGCTT 718
DB 624 TCTAGGGGGGTTCCCTACGGTT 646

RESULT 13

ADK46752

ID ADK46752 strand: cDNA; 826 BP.

XX ADK46752;

XX 21-APR-2005 (first entry)

XX Plant full length insert polynucleotide seqid 21492.

PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
PT pests, for conferring increased resistance to plant disease, or for
PT improving yield.

PS Claim 1; SEQ ID NO 25691; 15pp; English.

CC The invention describes a recombinant DNA construct comprising a
CC polynucleotide consisting of a sequence encoding an amino acid sequence
CC available in electronic form from the US patent office at
CC ftp.segdata.uspto.gov/sequence.html?docID:2004034888. The polynucleotide
CC of the invention are also useful in physical arrays of molecules and as
CC plant breeding markers. The recombinant DNA construct is useful for
CC improving plant tolerance to cold, heat, drought, herbicides, extreme
CC osmotic conditions, pathogens or pests, for manipulating growth rate in
CC plant cells by modification of the cell cycle pathway, for conferring
CC increased resistance to plant disease, for producing Galactomanan,
CC liginin or plant growth regulators, for increasing the rate of homologous
CC recombination in plants, for improving yield by modification of
CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC or by providing improved plant growth and development under at least one
CC stress condition or for modifying seed oil or protein yield and/or
CC content. This sequence represents a plant full length insert
CC polynucleotide that can be used in the recombinant DNA construct of the
CC invention.

XX Sequence 841 BP; 177 A; 237 C; 252 G; 175 T; 0 U; 0 Other;

Query Match 12.7%; Score 101.8; DB 13; Length 841;

Best Local Similarity 53.5%; Pred. No. 6.4e-22;

Matches 301; Conservative 0; Mismatches 232; Indels 30; Gaps 3;

QY 156 AGGAGACGACCTCGCAGGCCCCCTGAGGGTTCAGATTATGTGCAGAAACCTTTGGCGCTT 215
DB 113 AGGAGGCTGGGCGCAGGCCCCCGAGGAGCCATCTCTGCATCAATACCTGTGGCTTCT 172
QY 216 TCGGCGAGCCAGCTACCATGAGGGTGTGCTCGAAGTCTACCGAGAGACATGCAAG 275
DB 173 TCGGCGAGCCCGCGACATGAACTGTCTCCAGAGTCCCAAGAGAGATGATGCAAGC 232
QY 276 CGAAGATGACGGCTTAACTGTCAGCAAGCCTCAGGCTGCTCAGGCGACATGTCACAG 335
DB 223 AGGATCAGGCGCAGCTGCTGC-----CTCTCTATGTCAGACATGCTGTA 278
QY 336 CTGCTGCTGTTAAGCCCCCGCTCCTCTATGTCAGAGCCAGCTCAGAGGTTGAGA 395
DB 279 CGGCGAGCGACGCT-----CATGAGCCGCTGTTGCTGGCAGCAACG 324
QY 396 GAAACATGATTGTGCGCATCATCTTCAGCTATCAACAAGACCTGGTACCCCGCTG 455
DB 325 GTAGTAGCTGTGGCCCAAGTTGAGTTGCAAAATGAAAC--GTGACGACGCCCGCTGATG 382
QY 456 CAGCTGCCCTCAGGCAAGTAAGTCTCTATGTCAGAGCTCCCTCTAGAACCCGAGCCCAATC 515
DB 353 TTGCGCGAACCCAGCGAGGGGGGTGGGGGATCTCCAAAGGGGGGAAGTAGGGCCGAAAC 442
QY 516 GATGCGGATCTTGCAGAAAGCTGTGATTTGAAGAGATTAAAGTGTGCTGTGCAACC 575
DB 443 GGTGACGGGCTCGCAGAGAAAGGGTTGAGCTTACGGGATTTCAACTGCCGTGTGGAACT 502
QY 576 TCTACTGCGCTTAACTGCTGTAAGTCTGCGAACAACACTTGCACATATGACTCAAAAGCCG 635
DB 503 TGTATCTGGCACTCAACCGCTACTCTCGAAGACAGACGATGCAAGTTCCGACTACCGGACTG 562
QY 636 CAGGCGAGAGACGATTTGCGAAAGCTAATCTCTTGTGCTGCGCAGAGAGGTTGTCAAGT 695
DB 563 CTGCGCAGGAGCGCATTTGCGCAAGGCTAATCCGATGTGAAGGACAGCAAGCTCGACAGA 622
QY 636 TTTGATGAGCATCCGTTAAGCTT 718
DB 623 TCTAGGGGGGGTTCCTTACGGTT 645

RESULT 15

ADX50987
ID ADX50987 standard; cDNA; 846 BP.

AC ADX50987;

DT 21-APR-2005 (first entry)

DE Plant full length insert polynucleotide seqid 25727.

XX plant protectant; plant growth regulant; gene therapy; plant
XX recombinant DNA construct; physical array; plant breeding;
XX cold tolerance; heat tolerance; drought tolerance; herbicide
XX extreme osmotic condition; pathogen tolerance; pest tolerance
XX growth rate; cell cycle pathway; disease resistance;
XX galactomanan production; liginin production; plant growth
XX yield; plant growth; plant development; seed oil; protein
XX protein content; gene; ss.

OS Unidentified.

PN US2004034888-A1.

XX 19-FEB-2004.

XX 28-APR-2003; 2003US-00425114.

XX 06-MAY-1999; 99US-00304517.

XX 05-NOV-2001; 2001US-00985678.

PA (LILU/) LILU J.

PA (ZHOV/) ZHOV Y.

PA (KOVA/) KOVALIC D K.

PA (SCRE/) SCREEN S E.

PA (TABAS/) TABASRA J E.

PA (CAOV/) CAO Y.

PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao

XX WPI; 2004-180133/17.

XX New recombinant DNA construct, useful for improving plant

PT cold, heat, drought, herbicides, extreme osmotic condition

PT pests, for conferring increased resistance to plant disease

PT improving yield.

PS Claim 1; SEQ ID NO 25727; 15pp; English.

CC The invention describes a recombinant DNA construct comprising a
CC polynucleotide consisting of a sequence encoding an amino
CC available in electronic form from the US patent office at
CC ftp.segdata.uspto.gov/sequence.html?docID:2004034888. The
CC of the invention are also useful in physical arrays of mol
CC plant breeding markers. The recombinant DNA construct is u
CC improving plant tolerance to cold, heat, drought, herbicide
CC osmotic conditions, pathogens or pests, for manipulating g
CC plant cells by modification of the cell cycle pathway, for
CC increased resistance to plant disease, for producing galac
CC liginin or plant growth regulators, for increasing the rate
CC recombination in plants, for improving yield by modificati
CC photosynthesis or carbohydrate, nitrogen or phosphorus use
CC or by providing improved plant growth and development unde
CC stress condition or for modifying seed oil or protein yiel
CC content. This sequence represents a plant full length inse
CC polynucleotide that can be used in the recombinant DNA con
CC invention.

XX Sequence 846 BP; 178 A; 241 C; 252 G; 175 T; 0 U; 0 Other;

Query Match 12.7%; Score 101.8; DB 13; Length

Best Local Similarity 53.5%; Pred. No. 6.4e-22;

Matches 301; Conservative 0; Mismatches 232; Indels

QY 156 AGGAGACGACCTCGCAGGCCCCCTGAGGGTTCAGATTATGTGCAGAAACCT

215


```
Db 118 AGGAGGCTGGGTGGCCAGGCCCCCGAGGAGCCATCCTCTGCAATCAATACTGTGGCTTCT 177
Oy 216 TCGGCAAGCCAAAGCTAACATGGGGTTGTGCTGAAAGTGTACCGAGAGACAGTCATGCAAG 275
Db 178 TCGGCAGCGCGCGCAGCATGAAACATGTGCTCCAAAGTGCACAGAGAGATGATTAACGAAGC 237
Oy 276 CGAAGATGACGGCTTTAAGTGAAGCAAGCACTCAGGCTGCTCAGGCGAATCTGCCACAG 335
Db 238 AGGATCAGGCCCAAGCTGGCTGC-----CTCCTTATCGAACAGCATGTGAA 283
Oy 336 CTGCTGTGTTTCAGCCCCCGCTCCTGTACATGAGACCAAGCTCACATGCGAGTTGAGA 395
Db 284 CGGCAAGCAACGCGT-----CANGAGCCGGTTGTGTGCTGGCAGCAACAG 329
Oy 396 GAACATGATTTGTGCCGATCAATCTTCAGCTATCAACAAGACCTGTACCCCGCTG 455
Db 330 GTAGTAGCTGTGGCCCAAGTTGAGTGCACAAATGAAC--GTGCAGCAGCCCGCTGATG 387
Oy 456 CAGCTGCCCTCAGGCAATGAAGTCTGTATGCGAGCTCCCTTAGACCCGAGCCCAATC 515
Db 388 TTGCCGACCCAGCGAGGGGTGGCGGATCTCCAAAGGGGGAAGGTAGGCCGAAC 447
Oy 516 GATCGGATCTTGCAGAAAGCGTGTGATGACAGATTAAAGTGTGCTGTGCAAC 575
Db 448 GGTGACGGCCTGCAGAAAGGTTGGACTTAGGGAATTCACCTCCGGTGTGGAACT 507
Oy 576 TCTACTGCGCTTAACTGGTACTGACAAACACACTTGCACATATGACTACAAAGCG 635
Db 508 TGTACTGCGCACTCCACCGCTACTCCGACAGCAGACTGCAAGTTGACTACCGACTG 567
Oy 636 CAGGCGAGAAAGCGATTGGAAAGCTAATCCTTGTGTGCGCGAGAAAGTTGTCAAGT 695
Db 568 CTGCCAGGAGCGCCATGTGCAGAGCTAATCCGTTGTGAAGCAGACAAAGCTCGACAAAG 627
Oy 696 TTTGATGAGCATCCGTTAAGCTT 718
Db 628 TCTAGGGGGGTTCCCTACGGTT 650
```

Search completed: December 8, 2005, 07:58:35
Job time : 556 secs

Db	409	CCTCTACTGCGCTTTACATCGGTA	CTCGAACAACCACTTGCA	CATATGACTACCAAMGC	350
Qy	634	CGCAGGCGAGGAGGATTTGCGAAG	AGTAATCTCTGTTCGTGGCGAG	AGAGGTTGTCA	693
Db	349	CGCAGGCGAGGAGGATTTGCGAAG	AGTAATCTCTGTTCGTGGCGAG	AGAGGTTGTCA	290
Qy	694	GTTTGTGATGAGCATCCGTTAAG	CTTTTCTGCGCAGATTAGGCTT	CATACATTGAGTAA	753
Db	289	GTTTGTGATGAGCATCCGTTAAG	CTTTTCTGCGCAGATTAGGCTT	CATACATTGAGTAA	230
Qy	754	CTTACATCTTTCTCTTATTCGAG	AGAGCGGATCGCATCAAGA	797	
Db	229	CTTACATCTTTCTCTTATTCGAG	AGAGCGGATCGCATCAAGA	186	
RESULT 3					
LOCUS	BJS92791/c				
DEFINITION	BJS92791	733 bp	mRNA	linear	EST 22-OCT-2007
ACCESSION	BJS92791				
VERSION	BJS92791				
KEYWORDS	BJS92791.1	GI:37834779			
SOURCE	Physcomitrella patens subsp. patens				
ORGANISM	Physcomitrella patens subsp. patens				
REFERENCE	Eukaryota, Viridiplantae; Streptophyta; Embryophyta; Bryophyta; Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella.				
AUTHORS	Nishiyama, T., Fujita, T., Shin-i, T., Seki, M., Nishide, H., Uchiyama, I., Kamuya, A., Carninci, P., Hayashizaki, Y., Shinozaki, K., Kohara, Y. and Hasebe, M.				
TITLE	Comparative genomics of Physcomitrella patens gametophytic transcriptome and Arabidopsis thaliana: implication for land plant evolution				
JOURNAL	Proc. Natl. Acad. Sci. U.S.A.	100. (13)			8007-8012 (2003)
PUBMED	12808149				
COMMENT	Contact: Tadashi Shin-i Center For Genetic Resource Information National Institute of Genetics 111 Yata, Miehima, Shizuoka 411-8540, Japan Tel: 81-559-81-6856 Fax: 81-559-81-6855 Email: tshini@gemms.nig.ac.jp A backbone of the vector is basically from pBluescript II (KS), that was in vivo excised from a 1-F1C phage vector (Carninci et al 2001). 5' end of the cDNA that was digested with XhoI was ligated to SalI site of the vector and the 3' end including polyA tail was ligated to BamHI site of the vector (5'- GAGAGAGAGATCCAAACCTTGAGAGTATTTTATTTTATTTTAA-3' was used as a 1st 3' primer, and 5'-gggttcgagatcagctgtgtccagagagatgactcgagAACGNNNN-3' as 2nd 5'-hairpin primer, giving the following 5' boarder sequence, AGGCCAATCGAGCGAGCTGATGATCTCGAAGCC). cDNA insert could be amplified with conventional T7 and T3 primers. This full-length cDNA library was generated according to the method described in Nishiyama et al. (2003). Protoplasts were blended by the POLYTRON, and then cultivated on the BCDAYS medium for 13-14 days under the continuous light. These clones are available from Riken Bio Resource Center (http://www.brc.riken.go.jp/lab/ripd/Eng/index.html). The database of Physcomitrella EST clones is available at the PHYSCODbase (http://mes.nibb.ac.jp/). Location/Qualifiers 1..733 /organism="Physcomitrella patens subsp. patens" /mol_type="mRNA" /sub_species="patens" /db_xref="taxon:145401" /clone="npbh6h21" /issue_type="mixture of chloronemata, caulonemata and malformated buds" /clone_1fb="normalized full length cDNA library".				

ORIGIN	chloronemata, caulonemata and malformed but	
Query Match	72.4%; Score 582; DB 3; Length 73;	
Best Local Similarity	100.0%; Pred. No. 3e-168;	
Matches 582; Conservative 0; Mismatches 0; Indels 0		
Db	216 TCGCAGACCAAGCTACCAATGGAGTGTCTCGAAGTGTCTACCGAGAGAC	275
Db	723 TCGGAGGCCAAGCTACCAATGGAGTGTCTCGAAGTGTCTACCGAGAGAC	674
Qy	216 CGAAGATGACGGCTTTAGCTGAGCAAGCCACTCAGCTGTCCGAGCGAC	335
Db	673 CGAAGATGACGGCTTTAGCTGAGCAAGCCACTCAGCTGTCCGAGCGAC	614
Qy	336 CTGCTGCTGTTACAGCCCCCGCTCCTGTACATGATGAGACCAAGCTGACAT	395
Db	613 CTGCTGCTGTTACAGCCCCCGCTCCTGTACATGATGAGACCAAGCTGACAT	554
Qy	396 GAACAATGATTGTGCGCGCATCATCTTCCACATATCAACAAGACCTGCG	455
Db	553 GAACAATGATTGTGCGCGCATCATCTTCCACATATCAACAAGACCTGCG	494
Qy	456 CAGCTGCCCTCAGGCAATGAAATCTCTATCGCAGCTCCCTCTAGACC	515
Db	493 CAGCTGCCCTCAGGCAATGAAATCTCTATCGCAGCTCCCTCTAGACC	434
Qy	516 GATGCGGATCTTGCAGGAAAGCGTGTGATGATGACAGGATTTTAAGTGTG	575
Db	433 GATGCGGATCTTGCAGGAAAGCGTGTGATGATGACAGGATTTTAAGTGTG	374
Qy	576 TCTACTGCGCTTTACATCGGTACTCGACAAACACACTTGCACATATGA	635
Db	373 TCTACTGCGCTTTACATCGGTACTCGACAAACACACTTGCACATATGA	314
Qy	636 CAGGCGAGGAAAGCAATGCGAAAGCTAATCTTTGTCGTGCGCGAGAA	695
Db	313 CAGGCGAGGAAAGCAATGCGAAAGCTAATCTTTGTCGTGCGCGAGAA	254
Qy	696 TTGTATGAGCATCGGTAAAGCTTTTTCGCCACACATTTAGCTTCATAC	755
Db	253 TTGTATGAGCATCGGTAAAGCTTTTTCGCCACACATTTAGCTTCATAC	194
Qy	756 CTACATCTTTCTCTTTATCGAGAGAGGAGTGCATCAAGA 797	
Db	193 CTACATCTTTCTTTTATCGAGAGAGGAGTGCATCAAGA 152	
RESULT 4		
BU592532/c	729 bp	mRNA linear
LOCUS	BU592532	normalized full length cDNA library, chlor
DEFINITION	BU592532	caulonemata and malformed buds Physcomitrella patens
ACCESSION	BU592532	cDNA clone pphb6h03 3', mRNA sequence.
VERSION	BU592532.1	GI:37834520
KEYWORDS	EST.	
SOURCE	Physcomitrella patens subsp. patens	
ORGANISM	Physcomitrella patens subsp. patens	
REFERENCE	Buxaryota, Viridiplantae; Streptophyta; Embryophyta; Bryophyta; Funariidae; Funariales; Funariaceae; Phylloids (basses 1 to 729)	
AUTHORS	Nishiyama, T., Fujita, T., Shin-i, T., Seki, M., Nishiduchi, Y., Kamiya, A., Carninci, P., Hayashizaki, Y., Kohara, Y. and Hasebe, M.	K.,
TITLE	Comparative genomics of Physcomitrella patens gametophyte and Arabidopsis thaliana: implication	lant
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 100 (13), 8007-8012	
PIRMEED	12808149	
COMMENT	Contact: Tadashi Shin-i Center For Genetic Resource Information National Institute of Genetics	


```

OY 215 TTGGGAGCCAGCTACATAGGGGTTGTGCTCGAAGTGTACCGAGAGACAGTCATGCA 274
DB 747 TTGGGAGCCAGAGTACATAGGGGTTGTGCTCGAAGTGTACCGAGAGACAGTCATGCA 688
OY 275 GCGAAGATGACGGCTTTAGCTGAGCAAGCCACTCAGGCTGCTCAGGCGACATCTGCGACA 334
DB 687 GCGAAGATGACGGCTTTAGCTGAGCAAGCCACTCAGGCTGCTCAGGCGACATCTGCGACA 628
OY 335 GCTGCTGCTGTTAGAGCCCCCGCTCTGTTACATGAGACCAAGCTCACATGCGAGTTGAG 394
DB 627 GCTGCTGCTGTTAGAGCCCCCGCTCTGTTACATGAGACCAAGCTCACATGCGAGTTGAG 568
OY 395 AGAACAATGATTTGTGCGCGATCAATCTTCCAGCTATCAACAAGCTGGTATCCCGGCT 454
DB 567 AGAACAATGATTTGTGCGCGATCAATCTTCCAGCTATCAACAAGCTGGTATCCCGGCT 508
OY 455 GGAAGTCCCTTCAGGCAAGTGAAGTCTCTATGAGCACTCCCTTATGAGCCGAGCCCAAT 514
DB 507 GGAAGTCCCTTCAGGCAAGTGAAGTCTCTATGAGCACTCCCTTATGAGCCGAGCCCAAT 448
OY 515 CGATGCGGATCTTGGCGAGAAAGCTGTGATGATGACGATTTAAGTGTGCTGTGCGAAC 574
DB 447 CGATGCGGATCTTGGCGAGAAAGCTGTGATGATGACGATTTAAGTGTGCTGTGCGAAC 388
OY 575 CTCTACTGCGCTTTATACGATGCTGAGCAAAACACACTTGACATATGACTACAAAGCC 634
DB 387 CTCTACTGCGCTTTATACGATGCTGAGCAAAACACACTTGACATATGACTACAAAGCC 328
OY 635 GGAAGGCGAAGAAAGCAATTCGAAAGCTAACTCTTGTGCGCGAGAAAGTTGTCAAG 694
DB 327 GGAAGGCGAAGAAAGCAATTCGAAAGCTAACTCTTGTGCGCGAGAAAGTTGTCAAG 268
OY 695 TTTTGTATGACATCCGTTAAGCTTTTCTGCGAGATTTAAGCTTATACATTTAGTAAC 754
DB 267 TTTTGTATGACATCCGTTAAGCTTTTCTGCGAGATTTAAGCTTATACATTTAGTAAC 208
OY 755 TCTACATCTTCTCTTATCGAGAGGAGGATCGATCAAG 797
DB 207 TCTACATCTTCTCTTATCGAGAGGAGGATCGATCAAG 165

```

RESULT 6
 BJS505/c
 LOCUS
 DEFINITION
 BJS5055 727 bp mRNA linear EST 22-OCT-2003
 BJS5055 normalized full length cDNA library, chloronemata, cautionemata and malformed buds Physcomitrella patens subsp. patens cDNA clone pphb26n05 3', mRNA sequence.
 BJS5055
 BJS5055.1 GI:37827043
 EST.
 Physcomitrella patens subsp. patens
 Physcomitrella patens subsp. patens
 Bryophyta; Viridiplantae; Streptophyta; Embryophyta; Bryophyta; Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella. 1 (bases 1 to 727)
 Nishiyama, T., Fujita, T., Shin-1, T., Seki, M., Nishide, H., Uchiyama, I., Kamiya, A., Carninci, P., Hayashizaki, Y., Shinozaki, K., Kohara, Y. and Hasebe, M.
 Comparative genomics of Physcomitrella patens gametophytic transcriptome and Arabidopsis thaliana: implication for land plant evolution
 Proc. Natl. Acad. Sci. U.S.A. 100 (13), 8007-8012 (2003)
 12808149
 Contact: Tadaeu Shin-1
 Center For Genetic Resource Information
 National Institute of Genetics
 111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tshin@genes.nig.ac.jp
 A backbone of the vector is basically from pBluescript II (KS), that was in vivo excised from a 1-FLC phage vector (Carninci et al. 2001). 5' end of the cDNA that was digested with XhoI was ligated

to SalI site of the vector and the 3' end including ligated to BamHI site of the vector (5'- gagagagagagagatccaccctggagaggtttttttttttt used as a 1st 3' primer, and 5'-ggttctcgatgcacgctgttcacagacagcgatgactcgaaacggn1 5'-hairpin primer, giving the following 5' boarder AGGCCAATCGCCGACGATCGAATTCGCGAAGACCG). cDNA insert amplified with conventional T7 and T3 primers. This cDNA library was generated according to the method Nishiyama et al. (2003).
 Protonemata were blended by the POLYTRON, and then the BODATG medium for 13-14 days under the contin- These clones are available from RIKEN Bio Resource (http://www.brc.riken.go.jp/lab/epd/eng/index.html) of Physcomitrella EST clones is available at the PH (http://moses.nibb.ac.jp).

FEATURES
 source
 1..727
 /organism="Physcomitrella patens subsp. pa
 /mol_type="mRNA"
 /sub_species="patens"
 /db_xref="taxon:145481"
 /clone="pphb26n05"
 /issue_type="mixture of chloronemata, cau
 /clone_lib="normalized full length cDNA li
 chloronemata, cautionemata and malformed bu

ORIGIN
 Query Match 72.1%; Score 579.4; DB 3; Length 727
 Best Local Similarity 99.8%; Pred. No. 1.9e-167;
 Matches 580; Conservative 0; Mismatches 1; Indels 0;

```

OY 217 CGGACGCCAAGCTACATAGGGGTTGTGCTCGAAGTGTACCGAGAGAC 276
DB 727 CGGACGCCAAGCTACATAGGGGTTGTGCTCGAAGTGTACCGAGAGAC 668
OY 277 GAAAGTACGGCTTTAAGCTGAGCAAGCCACTCAGGCTGCTCAGGCGAC 336
DB 667 GAAAGTACGGCTTTAAGCTGAGCAAGCCACTCAGGCTGCTCAGGCGAC 608
OY 337 TGCTGCTGTTAGCCCCCGCTCTGTTACATGAGACCAAGCTCACATG 396
DB 607 TGCTGCTGTTAGCCCCCGCTCTGTTACATGAGACCAAGCTCACATG 548
OY 397 AACAAATGATTTGCGCGATCAATCTTCCAGCTATCAACAAGACTGTT 456
DB 547 AACAAATGATTTGCGCGATCAATCTTCCAGCTATCAACAAGACTGTT 488
OY 457 AGCTGCCCTTCAGGCAAGTGAAGTCTCTATGCGAGCTCCTTTAGACC 516
DB 487 AGCTGCCCTTCAGGCAAGTGAAGTCTCTATGCGAGCTCCTTTAGACC 428
OY 517 AAGCGATCTTGCAGAAAGCGTTGATGATGACAGATTTTAAGTGGC 576
DB 427 AATCGGATCTTGCAGAAAGCGTTGATGATGACAGATTTTAAGTGGC 368
OY 577 CTATGCGCTTTACATCGGTACTCGAGCAAAACACTTGCACATATGAC 636
DB 367 CTATGCGCTTTACATCGGTACTCGAGCAAAACACTTGCACATATGAC 308
OY 637 AGGCGAAGAGCGATTTGGAAGAACTAATCTTGTCTGTGCGCGAAG 696
DB 307 AGGCGAAGAGCGATTTGGAAGAACTAATCTTGTCTGTGCGCGAAG 248
OY 697 TTGATGAGATCCGTTAAGCTTTTCTGCGAGATTTAAGGCTTCTATCA 756
DB 247 TTGATGAGATCCGTTAAGCTTTTCTGCGAGATTTAAGGCTTCTATCA 188
OY 757 TACATCTTCTCTTATTCAGAGAGCGAGTGCATCAAGA 797
DB 187 TACATCTTCTCTTATTCAGAGAGCGAGTGCATCAAGA 147

```

RESULT 7
 BUS585014/c 721 bp mRNA linear EST 22-OCT-2003
 LOCUS BUS585014 normalized full length cDNA library, chloronemata, caulonemata and malformed buds Physcomitrella patens subsp. patens
 DEFINITION CDNA clone pphb26g04 3', mRNA sequence.
 ACCESSION BUS585014
 VERSION BUS585014.1 GI:37827002
 KEYWORDS EST.
 SOURCE Physcomitrella patens subsp. patens
 ORGANISM Physcomitrella patens subsp. patens
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta; Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella.
 REFERENCE 1 (bases 1 to 721)
 Nishiyama,T., Fujita,T., Shin-I,T., Seki,M., Nishide,H., Uchiyama,T., Kamiya,A., Carninci,P., Hayaishizaki,Y., Shinozaki,K., Kohara,Y. and Hasebe,M.
 Comparative genomes of Physcomitrella patens gametophytic transcriptome and Arabidopsis thaliana: implication for land plant evolution
 TITLE Proc. Natl. Acad. Sci. U.S.A. 100 (13), 8007-8012 (2003)
 JOURNAL 12808149
 PUBMED Contact: Tadao Shin-I
 COMMENT Center For Genetic Resource Information
 National Institute of Genetics
 1111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tshini@genes.nig.ac.jp
 A backbone of the vector is basically from pBluescript II (KS), that was in vivo excised from a 1-FLC phage vector (Carninci et al. 2001). 5' end of the cDNA that was digested with XhoI was ligated to SalI site of the vector and the 3' end including polyA tail was ligated to BamHI site of the vector(5'- gAAGAGAGAGATCCAAACCCCTGAGAGATTGTTTTTTTTTTVN-3' was used as a 1st 3' primer, and 5'-ggTTCGATGATCGCTGTCCAGACGATGATGACCGANNNN-3' as 2nd 5'-hairpin primer, giving the following 5' boarder sequence, AGGCCAAATCGCGCGAGCTGATTCGTCGAAACCG). cDNA insert could be amplified with conventional T7 and T3 primers. This full-length cDNA library was generated according to the method described in Nishiyama et al. (2003).
 Protonemata were blended by the POLYTRON, and then cultivated on the BCDATG medium for 13-14 days under the continuous light. These clones are available from RIKEN Bio Resource Center (http://www.brc.riken.go.jp/lab/epd/Eng/index.html). The database of Physcomitrella EST clones is available at the PHYSCOBASE (http://mos.nihb.ac.jp).
 FEATURES
 source location/Qualifiers
 1..721
 /organism="Physcomitrella patens subsp. patens"
 /mol_type="mRNA"
 /sub_species="patens"
 /db_xref="taxon:145481"
 /clone="pphb26g04"
 /cissue_type="mixture of chloronemata, caulonemata and malformed buds"
 /clone_1lb="normalized full length cDNA library, chloronemata, caulonemata and malformed buds"
 ORIGIN
 Query Match 70.9%; Score 570; DB 3; Length 721;
 Best Local Similarity 99.7%; Pred. No. 1,5e-164;
 Matches 570; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 226 AGCTACCATGGGGTGTGCTCGAAGTGCTACCGAGAGACATGCAAGCAAGATGAC 285
 Db 721 ACTACACATGGGGTGTGCTCGAAGTGCTACCGAGAGACATGCAAGCAAGATGAC 662
 QY 286 GCGTTTAACTGAGCAAGCACTCAGGCTGCTCGAGCGACATCTGCCACAGCTGCTGCT 345
 Db 661 GCGTTTAACTGAGCAAGCACTCAGGCTGCTCGAGCGACATCTGCCACAGCTGCTGCT 602

QY 346 TCAGCCCCCGGCTCTCTGTACATGAGACCAAGCTCATCATGCGAGTTGAG; 405
 Db 601 TCAGCCCCCGGCTCTCTGTACATGAGACCAAGCTCATCATGCGAGTTGAG. 542
 QY 406 TGTGCGCATCAATCTTCACACTATCAACAAGACCTGTGTTACCCCGCT 465
 Db 541 TGTGCGCATCAATCTTCACACTATCAACAAGACCTGTGTTACCCCGCT 482
 QY 466 TCAGGCACTGAAGTCTTATGCGAGCTCCTCTAGACCCGAGCCCAT 525
 Db 481 TCAGGCACTGAAGTCTTATGCGAGCTCCTCTAGACCCGAGCCCAT 422
 QY 526 TTGACAGAAAGGCTGTGATTGACAGATTAAAGTGTGCTGTGCAAC 585
 Db 421 TTGACAGAAAGGCTGTGATTGACAGATTAAAGTGTGCTGTGCAAC 362
 QY 586 TTATCATCGTATCTCGACAAACACACTTGACATATGACTACAAAGCC 645
 Db 361 TTATCATCGTATCTCGACAAACACACTTGACATATGACTACAAAGCC 302
 QY 646 AGCGATTGCGAAAGTAAATCTCTTGTGTGTGCGCGAGAAAGTTGCAAG 705
 Db 301 AGCGATTGCGAAAGTAAATCTCTTGTGTGTGCGCGAGAAAGTTGCAAG 242
 QY 706 ATCCGTTAAGCTTTTCTGCGACGATTAGGCTTCATATGAGTTAAC 765
 Db 241 ATCCGTTAAGCTTTTCTGCGACGATTAGGCTTCATATGAGTTAAC 182
 QY 766 CTTCTTATCGAGAGAGCGAGTCCGATCAAGA 797
 Db 181 CTTCTTATCGAGAGAGCGAGTCCGATCAAGA 150
 RESULT 8
 BUS606134/c 702 bp mRNA linear
 LOCUS BUS606134 normalized full length cDNA library, chlor. caulonemata and rhizoid-like protonemata Physcomitrella subsp. patens cDNA clone pphn35f21 3', mRNA sequence
 DEFINITION BUS606134
 ACCESSION BUS606134 GI:37848126
 VERSION BUS606134
 KEYWORDS EST.
 SOURCE Physcomitrella patens subsp. patens
 ORGANISM Physcomitrella patens subsp. patens
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryopsida; Funariidae; Funariales; Funariaceae; Pny; 1 (bases 1 to 702)
 Nishiyama,T., Fujita,T., Shin-I,T., Seki,M., Nishide,H., Uchiyama,T., Kamiya,A., Carninci,P., Hayaishizaki,Y., Kohara,Y. and Hasebe,M.
 Comparative genomes of Physcomitrella patens gamet transcriptome and Arabidopsis thaliana: implication evolution
 TITLE Proc. Natl. Acad. Sci. U.S.A. 100 (13), 8007-8012 (2003)
 JOURNAL 12808149
 PUBMED Contact: Tadao Shin-I
 COMMENT Center For Genetic Resource Information
 National Institute of Genetics
 1111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tshini@genes.nig.ac.jp
 A backbone of the vector is basically from pBluescript II (KS), that was in vivo excised from a 1-FLC phage vector (Carninci et al. 2001). 5' end of the cDNA that was digested with XhoI was ligated to SalI site of the vector and the 3' end including polyA tail was ligated to BamHI site of the vector(5'- gAAGAGAGAGATCCAAACCCCTGAGAGATTGTTTTTTTTTTVN-3' was used as a 1st 3' primer, and 5'-ggTTCGATGATCGCTGTTCAGACGATGATGACCGANNNN-3' as 2nd 5'-hairpin primer, giving the following 5' boarder sequence, AGGCCAAATCGCGCGAGCTGATTCGTCGAAACCG). cDNA insert

amplified with conventional T7 and T3 primers. The full-length cDNA library was generated according to the method described in Nishiyama et al. (2003). Protonemata were blended by the POLYTRON, and then cultivated on the BCDATG medium for 13-14 days under the continuous light. These clones are available from RIKEN Bio Resource Center (<http://www.brc.riken.go.jp/lab/epd/Eng/index.html>). The database of Physcomitrella EST clones is available at the PHYSCODbase (<http://mos.nibb.ac.jp/>).

FEATURES

source

Location/Qualifiers

1..702

/organism="Physcomitrella patens subsp. patens"

/mol_type="mRNA"

/sub_species="patens"

/db_xref="taxon:145481"

/clone="pbn35f21"

/issue_type="mixture of chloronemata, caulonemata and rhizoid-like protonemata"

/clone_1lb="normalized full length cDNA library, chloronemata, caulonemata and rhizoid-like protonemata"

ORIGIN

Query Match 68.8%; Score 553; DB 3; Length 702;

Best Local Similarity 100.0%; Pred. No. 2.8e-159;

Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

245 TCGAAGTGTCTACCGAGAGACATGCAAGCAAGATGACGGCTTTAGCTAGCAAGCC 304
 702 TCGAAGTGTCTACCGAGAGACATGCAAGCAAGATGACGGCTTTAGCTAGCAAGCC 643
 305 ACTCAGGCTGTCTACCGAGAGACATGCAAGCAAGATGACGGCTTTAGCTAGCAAGCC 364
 642 ACTCAGGCTGTCTACCGAGAGACATGCAAGCAAGATGACGGCTTTAGCTAGCAAGCC 583
 365 CATGAGACCAAGCTCATATGCGAGTTGAGAGAACATGATTTGCGCATCATCTTCC 424
 582 CATGAGACCAAGCTCATATGCGAGTTGAGAGAACATGATTTGCGCATCATCTTCC 523
 425 AGCTATCAACAAAGACTGTGTTACCCCGCTGCAAGCTGCGCAAGTAACTCTT 484
 522 AGCTATCAACAAAGACTGTGTTACCCCGCTGCAAGCTGCGCAAGTAACTCTT 463
 485 ATCGGAGCTCCCTCTGAGCCGAGCCCAATGATGCGAGTGTGCAAGGCTGTGA 544
 462 ATCGGAGCTCCCTCTGAGCCGAGCCCAATGATGCGAGTGTGCAAGGCTGTGA 403
 545 TTGACAGATTAAAGTGTGCTGTGCAACCTCTACTGCGCTTTACATCGTACTCGAC 604
 402 TTGACAGATTAAAGTGTGCTGTGCAACCTCTACTGCGCTTTACATCGTACTCGAC 343
 605 AAACACACTTGGACATATGATCAAAAGCGCGAGGCGAGAAAGCGATTGCGAAAGCTAAT 664
 342 AAACACACTTGGACATATGATCAAAAGCGCGAGGCGAGAAAGCGATTGCGAAAGCTAAT 283
 665 CCTCTTGTGTCGCGAGAGAGGTGTCAGTTTGTGAGCATCCGTTAAAGCTTTTTCG 724
 282 CCTCTTGTGTCGCGAGAGAGGTGTCAGTTTGTGAGCATCCGTTAAAGCTTTTTCG 223
 725 CGACGATTAAAGCTTCATCATTTAGTAAGTCTACATCTTTCTTTTATCGAGAGAGG 784
 222 CGACGATTAAAGCTTCATCATTTAGTAAGTCTACATCTTTCTTTTATCGAGAGAGG 163
 785 AGTCGATCAAGA 797
 162 AGTCGATCAAGA 150

RESULT 9

BJS91761/c

LOCUS BJS91761 704 bp mRNA linear EST 22-OCT-2003

DEFINITION BJS91761 normalized full length cDNA library, chloronemata, caulonemata and malformed buds Physcomitrella patens subsp. patens cDNA clone pbnh44121 3', mRNA sequence.

ACCESSION BJS91761
 VERSION BJS91761.1 GI:37833749

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

COMMENT

CONTACT

ADDRESS

TELEPHONE

FAX

EMAIL

ABSTRACT

NOTES

REFERENCES

COMMENTS

FEATURES

source

Location/Qualifiers

1..704

/organism="Physcomitrella patens subsp. p.

/mol_type="mRNA"

/sub_species="patens"

/db_xref="taxon:145481"

/clone="pbnh44121"

/issue_type="mixture of chloronemata, cau

malformed buds"

/clone_1lb="normalized full length cDNA 1;

chloronemata, caulonemata and malformed bu

ORIGIN

Query Match

Best Local Similarity 99.8%; Score 551; DB 3; Length 70;

Matches 551; Conservative 0; Mismatches 1; Indels

246 CGAAGTGTCTACCGAGAGACATGCAAGCAAGATGACGGCTTTAGCTAGCAAGCC 305
 704 TCGAAGTGTCTACCGAGAGACATGCAAGCAAGATGACGGCTTTAGCTAGCAAGCC 643
 306 ACTCAGGCTGTCTACCGAGAGACATGCAAGCAAGATGACGGCTTTAGCTAGCAAGCC 365
 644 ACTCAGGCTGTCTACCGAGAGACATGCAAGCAAGATGACGGCTTTAGCTAGCAAGCC 583
 366 CATGAGACCAAGCTCATATGCGAGTTGAGAGAACATGATTTGCGCATCATCTTCC 425
 584 CATGAGACCAAGCTCATATGCGAGTTGAGAGAACATGATTTGCGCATCATCTTCC 523
 426 AGCTATCAACAAAGACTGTGTTACCCCGCTGCAAGCTGCGCAAGTAACTCTT 485

Db 524 GGTATCAACAGACCTGTTACCCCGCTGACGTGCCCTCAGGCACTGATCTCTCTA 465
 QY 486 TCGAGCTCCCTCTGACCCCGAGCCCAATCGATTCGGATCTTGAGAGCGTTGTGAT 545
 Db 464 TCGAGCTCCCTCTGACCCCGAGCCCAATCGATTCGGATCTTGAGAGCGTTGTGAT 405
 QY 546 TGACAGATTTTAAAGTCTGCTGCGCAACCTTACTGCGCTTTTACATCGGTACTCGACA 605
 Db 404 TGACAGATTTTAAAGTCTGCTGCGCAACCTTACTGCGCTTTTACATCGGTACTCGACA 345
 QY 606 AACACACTTGCACATATGACTCMAAGCCGCGAGGCGAGAGACGATTCGGAAGCTAATC 665
 Db 344 AACACACTTGCACATATGACTCMAAGCCGCGAGGCGAGAGACGATTCGGAAGCTAATC 285
 QY 666 CTCTGTGTGCGCCGAGAAAGTTGCAAGTTTGATGAGACCTCGTTAGCTTTCTGTC 725
 Db 284 CTCTGTGTGCGCCGAGAAAGTTGCAAGTTTGATGAGACCTCGTTAGCTTTCTGTC 225
 QY 726 GACGATTTAGGCTTCATACATGAGTAACCTTCTTCTTTTATCGAGAGAGCA 785
 Db 224 GACGATTTAGGCTTCATACATGAGTAACCTTCTTCTTTTATCGAGAGAGCA 165
 QY 786 GTGCGATCAAGA 797
 Db 164 GTGCGATCAAGA 153

RESULT 10
 BUI65418/c 664 bp mRNA linear EST 16-OCT-2003
 LOCUS BUI65418 full length cDNA library, chloronemata and young
 DEFINITION gametophores Physcomitrella patens subsp. patens cDNA clone pph4c11
 3', mRNA sequence.
 BUI65418 GI:18333403
 BUI65418.1 GI:18333403

ACCESSION BUI65418.1 GI:18333403
 VERSION BUI65418.1
 KEYWORDS EST.
 SOURCE Physcomitrella patens subsp. patens
 ORGANISM Physcomitrella patens subsp. patens
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
 Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella.
 REFERENCE 1 (bases 1 to 664)
 AUTHORS Nishiyama,T., Fujita,T., Shin-I,T., Seki,M., Nishide,H.,
 Uchiyama,T., Kamiya,A., Carninci,P., Hayashizaki,Y., Shinzaki,K.,
 Kohara,Y. and Hasebe,M.
 Comparative genomics of Physcomitrella patens gametophytic
 transcriptome and Arabidopsis thaliana: implication for land plant
 evolution
 Proc. Natl. Acad. Sci. U.S.A. 100 (13), 8007-8012 (2003)
 JOURNAL 12808149
 PUBMED Contact: Tadao Shin-I
 COMMENT Center For Genetic Resource Information
 National Institute of Genetics
 1111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tshin@genes.nig.ac.jp

A backbone of the vector is basically from pbluescript II (KS),
 that was in vivo excised from a l-PIC phage vector (Carninci et al.
 2001). 5' end of the cDNA that was digested with XhoI was ligated
 to SalI site of the vector and the 3' end including polyA tail was
 ligated to BamHI site of the
 vector(5'- GAAGAAGAGAGATCCACCTGAGAGATTTTATTTTATTTTVA-3' was
 used as a 1st 3' primer, and
 5'-ggtctcgtgagtcacgtcgttcgcagacgcgatactgaaacggnnnn-3' as 2nd
 5'-hairpin primer, giving the following 5' boarder sequence,
 AGGCGCAATCGCGAGCTCGAATTCGCGAACCg). cDNA insert could be
 amplified with conventional T7 and T3 primers. This full-length
 cDNA library was generated according to the method described in
 Nishiyama et al. (2003).
 Protonemata were blended by the POLYTRON, and then cultivated on
 the BCDATG medium for 13-14 days under the continuous light.
 These clones are available from RIKEN Bio Resource Center
 (http://www.brc.riken.go.jp/lab/epd/Eng/index.html). The database

of Physcomitrella EST clones is available at the PF
 (http://mos.nibb.ac.jp).
 FEATURES
 source location/Qualifiers
 1. 664
 /organism="Physcomitrella patens subsp. pa
 /mol_type="mRNA"
 /sub_species="patens"
 /db_xref="taxon:145481"
 /clone="pph4c11"
 /tissue_type="mixture of chloronemata and
 gametophores with 2 to 5 leaves"
 /clone_id="full length cDNA library, chl
 young gametophores"

ORIGIN

Query Match 61.0%; Score 490.8; DB 3; Length 67
 Best Local Similarity 99.6%; Pred. No. 4.8e-140;
 Matches 492; Conservative 0; Mismatches 2; Indels 0;

QY 304 CACTCAGGCTGCTGAGGAGCATCTGCCAGACGCTGCTGTTACGCC 363
 Db 664 CACTCAGGCTGCTGAGGAGCATCTGCCAGACGCTGCTGTTACGCC 605
 QY 364 ACATGAGACCAAGCTCAGATCGAGTTGAGAGAACATGATTTGCCG 423
 Db 604 ACATGAGACCAAGCTCAGATCGAGTTGAGAGAACATGATTTGCCG 545
 QY 424 CAGCTATCAACAAGACCTGTTACCCCGCTGCGAGTCCCTCAGGCA 483
 Db 544 CAGCTATCAACAAGACCTGTTACCCCGCTGCGAGTCCCTCAGGCA 485
 QY 484 TATGCGAGCTCCCTCTAGACCCGAGCCCAATCGATCGGATCTTGCAGG 543
 Db 484 TATGCGAGCTCCCTCTAGACCCGAGCCCAATCGATCGGATCTTGCAGG 425
 QY 544 ATTGACAGATTTAAGTGTGCTGTGGCAACCTTACTGCGCTTTACAT 603
 Db 424 ATTGACAGATTTAAGTGTGCTGTGGCAACCTTACTGCGCTTTACAT 365
 QY 604 CMAACACCTTGACATATGACTACMAAGCCGCGAGGCGAGAGGATTT 663
 Db 364 CMAACACCTTGACATATGACTACMAAGCCGCGAGGCGAGAGGATTT 305
 QY 664 TCCTCTGTGTCGTCGCGAGAGGTTGTCAGTTTGTATGAGCATCGTT 723
 Db 304 TCCTCTGTGTCGTCGCGAGAGGTTGTCAGTTTGTATGAGCATCGTT 245
 QY 724 CCGAGATTTAGGCTTCATACATGAGTAACCTTACTTCTTCTTT 783
 Db 244 CCGAGATTTAGGCTTCATACATGAGTAACCTTACTTCTTCTTT 185
 QY 784 GAGTCGATCAAGA 797
 Db 184 GAGTCGATCAAGA 171

RESULT 11

BUI602584/c 642 bp mRNA linear EST 2003
 LOCUS BUI602584 normalized full length cDNA library, chloro
 DEFINITION caulonemata and rhizoid-like protonemata Physcomit
 subsp. patens cDNA clone pphba13 3', mRNA sequence

ACCESSION BUI602584.1 GI:37844576
 VERSION BUI602584.1
 KEYWORDS EST.
 SOURCE Physcomitrella patens subsp. patens
 ORGANISM Physcomitrella patens subsp. patens
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta;
 Bryopsida; Funariidae; Funariales; Funariaceae; Pny

REFERENCE 1 (bases 1 to 642)
 AUTHORS Nishiyama,T., Fujita,T., Shin-I,T., Seki,M., Nishide
 Uchiyama,T., Kamiya,A., Carninci,P., Hayashizaki,Y.,
 Kohara,Y. and Hasebe,M.

TITLE
Comparative genomics of *Physcomitrella patens* gametophytic transcriptome and Arabidopsis thaliana: implication for land plant evolution

JOURNAL
Proc. Natl. Acad. Sci. U.S.A. 100 (13), 8007-8012 (2003)

PUBMED
12808149

COMMENT
Contact: Tadaeu Shin-1
Center For Genetic Resource Information
National Institute of Genetics
111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshin@genes.nig.ac.jp
A backbone of the vector is basically from pBluescript II (KS), that was in vivo excised from a 1- ϕ C phage vector (Garnicki et al. 2001). 5' end of the cDNA that was digested with XhoI was ligated to SalI site of the vector and the 3' end including polyA tail was ligated to BamHI site of the vector (5'-gagagagagagatccacacccctgagagattttttttttttttt-3' was used as a 1st 3' primer, and 5'-ggttctgagtcagtcgttccagacagcgatgactgagaaacggnnnn-3' as 2nd 5'-hairpin primer, giving the following 5' border sequence, AGGCCAATCGCCGAGCTCGAATTCGTGAGAACCG). cDNA insert could be amplified with conventional T7 and T3 primers. This full-length cDNA library was generated according to the method described in Nishiyama et al. (2003).
Protonemata were blended by the POLYTRON, and then cultivated on the BCDATG medium for 13-14 days under the continuous light. These clones are available from RIKEN Bio Resource Center (<http://www.brc.riken.go.jp/lab/epd/Eng/Index.html>). The database of *Physcomitrella* EST clones is available at the PHYSCODbase (<http://moss.nibb.ac.jp>).
location/Qualifiers
1. 642
/organism="Physcomitrella patens subsp. patens"
/mol_type="mRNA"
/sub_species="patens"
/db_xref="taxon:145481"
/clone="pPhn8a13"
/tissue_type="mixture of chloronemata, caulonemata and rhizoid-like protonemata"
/clone_1lb="normalized full length cDNA library, chloronemata, caulonemata and rhizoid-like protonemata"

ORIGIN
Query Match 56.1%; Score 451.4; DB 3; Length 642;
Best Local Similarity 99.3%; Pred. No. 7,56-128;
Matches 452; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

343 TGTTCAGCCCCCGCTCTGTACATGAGACCAAGCTCAGCGAGTTGAGAGAACAT 402
|||
642 TGTTCAGCCCCCGCTCTGTACATGAGACCAAGCTCAGCGAGTTGAGAGAACAT 583
|||
403 GATTGTGCGCATCAATCTTCCAGCTATCAACAGACCTGTTTACCCCGCTGCACTGC 462
|||
582 GATTGTGCGCATCAATCTTCCAGCTATCAACAGACCTGTTTACCCCGCTGCACTGC 523
|||
463 CCCTCAGCAGTGAAGTCTCTATGCAAGCTCTCTTGAAGCCGAGCCCAATGATGCGG 522
|||
522 CCNTCAGCAGTGAAGTCTCTATGCAAGCTCTCTTGAAGCCGAGCCCAATGATGCGG 463
|||
523 ATCTTGAGAGGAGGTGTTGATTGACAGATTTAAGTGTGCTGTGCAACCTTACTG 582
|||
462 ATCTTGAGAGGAGGTGTTGATTGACAGATTTAAGTGTGCTGTGCAACCTTACTG 403
|||
583 CGCTTTACATCGGTAATCTGGAACAAACACATTTGACATGCTCAAAAGCGCAGGCA 642
|||
402 CGCTTTACATCGGTAATCTGGAACAAACACATTTGACATGCTCAAAAGCGCAGGCA 343
|||
643 GGAAGCGATTGCGAAAGCTAATCTCTTGTGTGCGCGGAGAAAGTTGTCAAGTTTGTATG 702
|||
342 GGAAGCGATTGCGAAAGCTAATCTCTTGTGTGCGCGGAGAAAGTTGTCAAGTTTGTATG 283
|||
703 AGCATCCGTTAAGCTTTCTGCGGAGATTAGGCTTCAATGAGTAATCTACATC 762
|||

DB 282 AGCATCCGTTAAGCTTTCTGCGGAGATTAGGCTTCAATGAGTAATCTACATC 223
|||

QY 763 TTTCTTCTTTATCGAGAGCGAGTGCATCAAGA 797
|||

DB 222 TTTCTTCTTTATCGAGAGCGAGTGCATCAAGA 188
|||

RESULT 12
BU941820 719 bp mRNA linear
LOCUS BU941820 pphf full-length cDNA library Physcomitrell
DEFINITION patens cDNA clone pphf16106 5', mRNA sequence.
ACCESSION BU941820
VERSION BU941820.1 GI:67568996
KEYWORDS EST.
SOURCE Physcomitrella patens subsp. patens
ORGANISM Physcomitrella patens subsp. patens
Bryopsida; Funariidae; Funariales; Funariaceae; Ph
1 (bases 1 to 719)
Fujita, T., Nishiyama, T., Shin-1, T., Kohara, Y. and N
Physcomitrella patens EST at a stage of the first a
division of protonemata
Unpublished (2005)
Contact: Tadaeu Shin-1
Center For Genetic Resource Information
National Institute of Genetics
111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshin@genes.nig.ac.jp
Protonemata were inoculated on BCDATG medium for ev
Protonemata were isolated from the protonemata, fun
at 25C under continuous light for 2-3 days. The reg
which were rich in cells at a stage during the first
cell division, were collected. Total RNA was extrac
constructing a full-length cDNA library. The databa
clones is available at the PHYSCODbase (<http://moss.nibb.ac.jp>).
location/Qualifiers
1. 719
/organism="Physcomitrella patens subsp. pa
/mol_type="mRNA"
/sub_species="patens"
/db_xref="taxon:145481"
/clone="pPhf16106"
/tissue_type="regenerated protonemata (ch1
/dev stage="at the first asymmetric cell d
protonemata"
/clone_1lb="pphf full-length cDNA library"
/note="Protonemata were inoculated on BCD
every ca. 5 days. Protonemata were isolate
protonemata, further incubated at 25C unde
light for 2-3 days. The regenerated cells,
in cells at a stage during the first asymm
division, were collected. Total RNA was ex
constructing a full-length cDNA library."

ORIGIN
Query Match 42.9%; Score 344.8; DB 3; Length 7
Best Local Similarity 72.7%; Pred. No. 7,66-95;
Matches 445; Conservative 0; Mismatches 167; Indels 0;

QY 162 CGACTCGCAGGCCCTGAGGGTCCAGTTATGTCACAAACCTTGGCGG 221
|||
DB 1 CAACCTCCACAGCCCGAGGGGCTGTGATGTGTGAAGACGTGTGCGG 60
|||
QY 222 GCCAAGCTACATGGGGTTGTGCTCGAAGTGTACCGAGACAGTCAT 281
|||
DB 61 GCCAAGCTACATGGGGTTGTGCTCGAAGTGTGTACAGGAGACCGTCAT 120
|||
QY 282 TGAGGGCTTTAGCTGAGCAAGCACTCAGGCTGCTCAGGCGCAATCTGC 341
|||

Db 121 TGACCGCCGTGGCCGAGCAGAGCCACTCAAGCCGCTCAGGTGCTGCATCCGCTGATCCA 180

Qy 342 CTGTTCAAGCCCCCGCTCTGTACATGAGACCACTACATGCGAGGTGAGAGAAACA 401

Db 181 GTGCTCACCTCCAGATTCTTAAGAGAGAGACAGAGCTCATTTGAGGCCGACAGCATGT 240

Qy 402 TGATTGTCGGCATCATCTTTCACGCTATCAACAAGACCTGGTTACCCCGCTGACGCTG 461

Db 241 TGATCAACCCGCTCAATCTTCCAGCCACATCCAGTTGAGGTTCTCCGGTTACAGTAG 300

Qy 462 CCCCTCAGGCAAGTAAGTCTCTATCGAGCTCCCTTAGACCCGACCCCAATGATGCG 521

Db 301 CCCCCAGAGTTGTAGTGGCCCCGTTGCAACACCTCCAGACCGCACCAACCGGTCGG 360

Qy 522 GATCTTGCAGAGAGCGTGTGATTGACAGAGATTAAAGTGTGCTGTGCAACCTCTACT 581

Db 361 GATCTTGCAGAGAGCGCGTGAGATTGACAGATTTCAGTGTGTTGGGACCTTTT 420

Qy 582 GCGCTTACATGCTGATCGACAAACACACTTGACATATGACTACAAAGCCGACGAGGC 641

Db 421 GTGCTTACATGCTGATCTGACAAACATAGTTGACATTATTAACAAGCTGCTGAC 480

Qy 642 AGGAAGCAATTCGAAAGTAATCTTGTGCTGCGGAGAGAGTTGCAAGTTTGAT 701

Db 481 AGGAGGCTATTGCAAGGCTAATCTTGTGCTGCGGAGAGAGTTGTAAGTTCTGAA 540

Qy 702 GAGCATCCGTTAAGCTTTTCTGCCAGCATTTAGGCTTATCATGACTGACTTACT 761

Db 541 GAGGAGCTGCTGACCTTTTGTGATGTGACGATGAGCTTTCAGCATGAAATGTTCTTAAC 600

Qy 762 CTTTCTCTTAA 773

Db 601 TTTCTTTATTGA 612

RESULT 13
 BJ605573/c 738 bp mRNA linear EST 22-OCT-2003
 LOCUS
 DEFINITION

ACCESSION BJ605573 normalised full length cDNA library, chloronemata, caulonemata and rhizoid-like protonemata Physcomitrella patens subsp. patens cDNA clone pphn34m03 3', mRNA sequence.

VERSION BJ605573.1 GI:37847565

KEYWORDS EST

SOURCE Physcomitrella patens subsp. patens

ORGANISM Physcomitrella patens subsp. patens Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta; Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella.

REFERENCE 1 (bases 1 to 738) Nishiyama, T., Fujita, T., Shin-I, T., Seki, M., Nishide, H., Uchiyama, I., Kamiya, A., Carninci, P., Hayashizaki, Y., Shinozaki, K., Kohara, Y. and Hasebe, M.

AUTHORS Comparative genomes of Physcomitrella patens gametophytic transcriptome and Arabidopsis thaliana: implication for land plant evolution

TITLE Proc. Natl. Acad. Sci. U.S.A. 100 (13), 8007-8012 (2003)

JOURNAL 12808149

PUBMED Contact: Tadao Shin-I

COMMENT Center For Genetic Resource Information

National Institute of Genetics

111 Yata, Mishima, Shizuoka 411-8540, Japan

Tel: 81-559-81-6856

Fax: 81-559-81-6855

Email: tshin@genes.nig.ac.jp

A backbone of the vector is basically from pBluescript II (KS), that was in vivo excised from a 1-F1C phage vector (Carninci et al. 2001). 5' end of the cDNA that was digested with XhoI was ligated to SalI site of the vector and the 3' end including polyA tail was ligated to BamHI site of the vector (5'-gagagagagagagatccacccttgagagagattttttttttttttt-3' was used as a 1st 3' primer, and 5'-gggttcgagatcgtcttcacacacgcatgactgagaaacgannnnn-3' as 2nd 5'-hairpin primer, giving the following 5' border sequence,

AGGCCAAATCGCCGAGCTCGAATTCGTGAGAACCG). cDNA insert amplified with conventional T7 and T3 primers. This cDNA library was generated according to the method Nishiyama et al. (2003). Protonemata were blended by the POLYTROX, and then the BCDATG medium for 13-14 days under the contin These clones are available from RIKEN BGI Resource (<http://www.brc.riken.go.jp/lab/epd/Eng/index.html>) of Physcomitrella EST clones is available at the PH (<http://mos.nibb.ac.jp>).

FEATURES

source
 1..738
 /organism="Physcomitrella patens subsp. pa"
 /sub_type="mRNA"
 /db_xref="taxon:145481"
 /clone="pph34m03"
 /tissue_type="mixture of chloronemata, cau rhizoid-like protonemata"
 /clone_1ib="normalized full length cDNA 1i chloronemata, caulonemata and rhizoid-like

ORIGIN

Query Match 41.6%; Score 334.4; DB 3; Length 7
 Best Local Similarity 72.3%; Pred. No. 1.3e-91;
 Matches 434; Conservative 0; Mismatches 166; Indels 0;

Qy 174 CCCCGAGGGTCCAGTTATGTCAGAAACCTTGGGGCTTCTTGGGCGAG 233

Db 738 CCCGAGGGGGCCCTGTGATGTGTAAAGCGTGTGGGTTCTTGGGCGAG 679

Qy 234 TGGGGTGTGCTCGAAGTGTACCGAGAGACATGATCAACGAGAGAT 293

Db 678 TGGGCTCTGCTCGAAGTGTACCGAGAGACATGATCAACGAGAGAT 619

Qy 294 CTGAGCAAGCCACTGAGGCTCTAGGCGGACATCTGCCAGCTGCTGC 353

Db 618 CCGAGCAGGCCACTCAAGCCGCTCAGGTGCTGCCATCCGCTGCATCCAG 559

Qy 354 CCGCTCTGTATGATGACCAAGCTCAGATGCGAGGTTGAGAGAACAT 413

Db 558 CAGTCTTATGAGAGAGCAAGAGCTAATTTGAGGCGGACAGCATGTT 499

Qy 414 ATCAATCTTCCAGCTATCAACAAAGCTGTATACCCCGCTGCAGCTGC 473

Db 498 CTCAATCTTCCAGCCACATCCAGTTGAGGTTGCTCCGTTACATAGC 439

Qy 474 TGAAGTCTCTATGCGACCTCTCTTAGACCCGAGCCCAATGATGCG 533

Db 438 TAGTTGCCCCCGTTGCCACACCTCCAGACCGGCACCAACCGGTGCG 379

Qy 534 AGCGTGTGATGACAGAGTTTAAGTGTGCTGTGGGCAACCTCTACAG 593

Db 378 AGCGGTGGGATGACAGAGATTTCAGTGTGTTGGGCAACCTTTT 319

Qy 594 GGTACTCGAACAACACACTTGCACATATGACTACAAAGCCGAGGCGA 653

Db 318 GGTATTTGACAAACATATGTTGCACTTATGATTACAAAGCTGTGACA 259

Qy 654 CGAAAGCTAATCTTGTGCGGCGGAGAGGTTGCAAGTTTGTATGAT 713

Db 258 CAAAGGCTAATCTTGTGTTGGGCGGAGAGTTGTTAAAGTTGGAAG 199

Qy 714 AGCTTTTGTGCGGAGATTTAGGCTTCTATCATTTAGATGATCAATC 773

Db 198 GACCTTTTGTGATGTGACGTGAGCTTCAGGATGGAATGATTTTAACT 139

RESULT 14
 BJ586048/c 736 bp mRNA linear
 LOCUS BJ586048 normalised full length cDNA library, chloronemata and malformed buds Physcomitrella patens
 DEFINITION

CDNA clone pphb29418 3', mRNA sequence.

ACCESSION BJ586048
 VERSION BJ586048.1 GI:37828036
 KEYWORDS EST.
 SOURCE Physcomitrella patens subsp. patens
 ORGANISM Physcomitrella patens subsp. patens
 Bryopsida; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
 Bryopsida; Funariaceae; Funariales; Funariaceae; Physcomitrella.
 1 (bases 1 to 736)
 AUTHORS Nishiyama,T., Fujita,T., Shin-I,T., Seki,M., Nishide,H.,
 Uchiyama,I., Kamiya,A., Carninci,P., Hayashizaki,Y., Shinzaki,K.,
 Kohara,Y. and Hasebe,M.
 TITLE Comparative genomics of Physcomitrella patens gametophytic
 transcriptome and Arabidopsis thaliana: implication for land plant
 evolution
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 100 (13), 8007-8012 (2003)
 PUBMED 12808149
 COMMENT Contact: Tadasu Shin-I
 National Institute of Genetics
 1111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tsuhin@genes.nig.ac.jp
 A backbone of the vector is basically from pBluescript II (KS),
 that was in vivo excised from a 1-FLC phase vector (Carninci et al.
 2001). 5' end of the cDNA that was digested with XhoI was ligated
 to SalI site of the vector and the 3' end including polyA tail was
 ligated to BamHI site of the
 vector(5'- gagaGAGAGAGATCCAAACCTGAGAGATTTTCTTTT-3' was
 used as a 1st 3' primer, and
 5'-ggTCTCGAGTCATCGCTTCCGACAGCGATGATGAAACGNNNN-3' as 2nd
 5'-hairpin primer, giving the following 5' boarder sequence,
 AGGCCAAATGCGCCGATCGATTCGATTCGGAACCG). cDNA insert could be
 amplified with conventional T7 and T3 primers. This full-length
 cDNA library was generated according to the method described in
 Nishiyama et al. (2003).
 Protonemata were blended by the POLYTRON, and then cultivated on
 the BCDATG medium for 13-14 days under the continuous light.
 These clones are available from RIKEN Bio Resource Center
 (http://www.brc.riken.go.jp/lab/epd/Eng/index.html). The database
 of Physcomitrella EST clones is available at the PHYSCobase
 (http://mes.nibb.ac.jp).
 FEATURES
 source location/Qualifiers
 1..736
 /organism="Physcomitrella patens subsp. patens"
 /mol_type="mRNA"
 /sub_species="patens"
 /db_xref="taxon:145481"
 /clone="pphb29418"
 /cisse_type="mixture of chloronemata, caulonemata and
 malformed buds"
 /clone_lib="normalized full length cDNA library,
 chloronemata, caulonemata and malformed buds"

ORIGIN

Query Match 41.1%; Score 330.4; DB 3; Length 736;
 Best Local Similarity 72.6%; Pred. No. 2.2e-90;
 Matches 427; Conservative 0; Mismatches 161; Indels 0; Gaps 0;

186 CAGTATATGCAAGAACTTGGGGCTTCTGGGACGCAAGCTACCAATGGGTTGTGCT 245
 DB CTGTATATGTAAGAACTGTGGCTTCTTGGGACGCAAGCTACCAATGGGTTGTGCT 677

246 CGAAGTGTACCGAGAGACAGTCATGCAAGCAAGATGACGCGCTTAACTGAGCAAGCA 305
 DB CGAAGTGTACCGAGAGACCGTCATGCAAGCAAGATGACGCGCTTAACTGAGCAAGCA 617

306 CTCAGGCTGCTAGGCGAGACATTTGCCACAGCTGCTGCTTTTCAAGCCCCGCTCTGAC 365
 DB CTCAGGCGGCTCAGGCTGCTGCTCCATCCGCTGATCCGCTGCTCAGGCTCTTAATAG 557

366 ATGAGACCAAGCTCATGCGAGGTTGAGAGAAATGATTTGTGCGGATCAATCTTCCA 425

DB 556 AAGAGACAGAGCTCATTTAGAGCCGACAGCATGTTGATTAACCCGCC
 QY 426 GCTATCAACMAAGCTGTGTTAACCCCGCGCTGACAGCTGCCCTCAGGACAGT
 DB 496 GCCACATCAAGTTGAGGTTGTCTCCGGTTACAGTATGCCCCCAGGTTG
 QY 486 TGGCAGCTCCCTCTAGACCCGAGCCCAATGATGGGATCTTGCAGGA,
 DB 436 TTGCCACACCCCTCCAGACCGGACCAACCCGCTGGGAGATCTTGACGAA,
 QY 546 TGACGAGATTTAAGTGTGCTGGGCAACCTTACATGCGCTTAAATGC
 DB 376 TGACGAGATTTCAAGTGTGCTGGGACCTTTTGTGCTTAAATGC
 QY 606 AACACATTTGACATATGACTACAAAGCCGACGAGGACAGAACGATTC
 DB 316 AACATAGTTGACATATGATTAACAGGCTGCTGACAGAGAGCTATTGC
 QY 666 CTCTTGTGCTGCGCGAAGAGTTTCAAGTTTGAATGACATCCGTTA,
 DB 256 CTTTATGTTGTGCGCGAAGAGTTTGAATGTTCTGAAGAGGACCTCTC,
 QY 726 GACGATTTAGGCTCATATGAGTAACTACATCTTCTCTTTA
 DB 196 GTGACGATAGGCTTACGCGATGAAATGTTCTTAACTTTCTTATTTGA

RESULT 15
 BJ598461/c
 LOCUS
 DEFINITION
 BJ598461 normalized full length cDNA library, chloronemata and rhizoid-like protonemata Physcomitrella patens subsp. patens
 accession BJ598461
 version BJ598461.1 GI:37840453
 keywords EST
 source Physcomitrella patens subsp. patens
 organism Physcomitrella patens subsp. patens
 Bryopsida; Viridiplantae; Streptophyta; Embryophyta;
 Bryopsida; Funariaceae; Funariales; Funariaceae; Physcomitrella.
 1 (bases 1 to 736)
 authors Nishiyama,T., Fujita,T., Shin-I,T., Seki,M., Nishide,H.,
 Uchiyama,I., Kamiya,A., Carninci,P., Hayashizaki,Y.,
 Kohara,Y. and Hasebe,M.
 title Comparative genomics of Physcomitrella patens gametophytic
 transcriptome and Arabidopsis thaliana: implication for land plant
 evolution
 journal Proc. Natl. Acad. Sci. U.S.A. 100 (13), 8007-8012
 pubmed 12808149
 comment Contact: Tadasu Shin-I
 National Institute of Genetics
 1111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tsuhin@genes.nig.ac.jp
 A backbone of the vector is basically from pBluescript II (KS),
 that was in vivo excised from a 1-FLC phase vector
 (Carninci et al. 2001). 5' end of the cDNA that was digested with XhoI
 was ligated to SalI site of the vector and the 3' end including
 polyA tail was ligated to BamHI site of the
 vector(5'- gagaGAGAGAGATCCAAACCTGAGAGATTTTCTTTT-3' was
 used as a 1st 3' primer, and
 5'-ggTCTCGAGTCATCGCTTCCGACAGCGATGATGAAACGNNNN-3' as 2nd
 5'-hairpin primer, giving the following 5' boarder sequence,
 AGGCCAAATGCGCCGATCGATTCGATTCGGAACCG). cDNA insert could be
 amplified with conventional T7 and T3 primers. This full-length
 cDNA library was generated according to the method described in
 Nishiyama et al. (2003).
 Protonemata were blended by the POLYTRON, and then cultivated on
 the BCDATG medium for 13-14 days under the continuous light.
 These clones are available from RIKEN Bio Resource

(<http://www.brc.riken.go.jp/lab/spd/Eng/index.html>). The database of *Physcomitrella* EST clones is available at the PHYCObase (<http://mos.nibb.ac.jp>).

FEATURES

SOURCE

Location/Qualifiers

1. . 723

/organism="Physcomitrella patens subsp. patens"

/mol_type="mRNA"

/sub_species="patens"

/db_xref="taxon:145481"

/clone="ppln23j11"

/tissue_type="mixture of chloronemata, caulonemata and rhizoid-like protonemata"

/clone_lib="normalized full length cDNA library,"

chloronemata, caulonemata and rhizoid-like protonemata"

ORIGIN

Query Match 41.0%; Score 329.8; DB 3; Length 723;
Best Local Similarity 72.5%; Pred. No. 3.4e-90;
Matches 427; Conservative 0; Mismatches 162; Indels 0; Gaps 0;

```
OY 185 CCAGTTATGCAAGAACCTTTGGCGCTTCTTCGCGACCAAGCTACCATGGGTTGTGC 244
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 723 CCTGATGATGTAAAGACGTGCGGTTCTTTGGCAGCCAGCCATATGGGTTCTGCG 664

OY 245 TCGAAGTCTACCGAGAGCAAGTCATGCAAGGCAAGATGACGGCTTTAGCTGAGCAAGCC 304
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 663 TCGAAGTCTTACAGGAGAACGTCATGCAAGGCAAGATGACGGCGGTGCGAGCAAGGCC 604

OY 305 ACTCAGGCTGCTCAGGCGACATCTGCAACGCTGCTGTTCAAGCCCCCGGCTCTGTA 364
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 603 ACTCAAGCCGCTCAGGTGCTGCATCCGTCGATCCAGTGCATCAGCTCCAGTTCTATG 544

OY 365 CATGAGACCAAGCTCAGATGCGAGGTTGAGAGAACATGATGTGCCCCGATCATCTCC 424
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 543 GAAGGAGCAAGAGCTCAATTGAGGCCGACAGCATGTTGATACAAACGGCTCAATCTTCC 484

OY 425 AGCTATCAACAAGACTGATACCCCGCTGAGCTGCCCTCAGGAGTGAAGTCTCT 484
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 483 AGCCACCATTCAGTTGAGGTTGCTCCGGTTACATAGCCCCCAGGTTGATGTCCTCC 424

OY 485 ATCGAGCTCCCTCTAGACCCGAGCCCAATCGATCGGATCTTGCAGGAAGCGTTGGA 544
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 423 GTTGCACACCCCTCCAGACCGGCACCAAAACGGTCCGATCTTGCAGGAAGCGGTGGA 364

OY 545 TTGACAGATTTAAGTGTGCTGTGCGAACCTCTACTGCGCTTTACATCGTACTCGAC 604
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 363 TTGACAGATTTCAAGTGTGCTGTGCGAACCTTTTGTGCTTTACATCGTACTCTGAC 304

OY 605 AAACACACTTGCACATATGACTACAAAGCCGAGGCGAGAGCGATTGGCAAGCTAAT 664
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 303 AAACATAGTTGCACTTATGATTAACAAGGTGCTGGAACGAGGCTATTTGCCAAAGCTAAT 244

OY 665 CCTTGTGCTGCGCGAGAAAGGTTGCAAGTTTGTATGAGCATCCGTTAAGCTTTTCTGC 724
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 243 CCTTTAGTTGTGCGCGAGAAAGGTTGTTAAGTTCTGAAGAGGAGCTGCTGACCTTTTGA 184

OY 725 CGACGATTTAGGCTTCATACATTTGAGTAAGTCTACATCTTTCTTTA 773
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 183 TGTGACGTAAGGCTTCAGGCAATGATAGTTCTTAACTTTCTTTATGTA 135
```

Search completed: December 8, 2005, 10:12:06
Job time : 3734 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using bw model

Run on: December 8, 2005, 07:38:03 ; Search time 199 Seconds
(without alignments)
7181.706 Million cell updates/sec

Title: US-10-716-089-12

Perfect score: 804
Sequence: 1 atccgcggcaccagctccgc.....agtcgcatcagaagctcgcc 804

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents NA:*

- 1: /cgn2_6/prodata/1/ina/1_COMB_seq.*
- 2: /cgn2_6/prodata/1/ina/5_COMB_seq.*
- 3: /cgn2_6/prodata/1/ina/6A_COMB_seq.*
- 4: /cgn2_6/prodata/1/ina/6B_COMB_seq.*
- 5: /cgn2_6/prodata/1/ina/7_COMB_seq.*
- 6: /cgn2_6/prodata/1/ina/8_COMB_seq.*
- 7: /cgn2_6/prodata/1/ina/9_COMB_seq.*
- 8: /cgn2_6/prodata/1/ina/10_COMB_seq.*
- 9: /cgn2_6/prodata/1/ina/11_COMB_seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	804	100.0	804	3	US-09-828-303-12
2	776.4	96.6	1531	3	US-09-828-303-4
3	96.8	12.0	283	3	US-09-313-294A-7446
4	93.6	11.6	281	3	US-09-313-294A-1951
5	78.8	9.8	211	3	US-09-313-294A-3856
6	77	9.6	1197	2	US-08-861-263-4
7	77	9.6	1197	2	US-09-134-596-4
8	77	9.6	1197	2	US-09-293-273-4
9	74.2	9.2	2425	3	US-09-949-016-810
10	74.2	9.2	2684	3	US-09-949-016-3693
11	71.8	8.9	2515	3	US-09-244-805-10
12	61	7.6	14524	3	US-09-949-016-1552
13	61	7.6	14785	3	US-09-949-016-15435
14	59.4	7.4	348	3	US-09-640-211A-1742
15	54.6	6.8	291	3	US-09-313-294A-4836
16	54.6	6.8	306	3	US-09-313-294A-6731
17	54.6	6.8	306	3	US-09-313-294A-5600
18	54.6	6.8	307	3	US-09-313-294A-6505
19	52	6.5	265	3	US-09-313-294A-3333
20	49.8	6.2	317	3	US-09-640-211A-1723
21	49.8	6.2	355	3	US-09-640-211A-1657
22	49.8	6.2	406	3	US-09-640-211A-1709
23	46.6	5.8	362	3	US-09-640-211A-1637
24	46.6	5.8	389	3	US-09-640-211A-1587

25	46.2	5.7	329	3	US-09-640-211A-1702	Seq
26	44	5.5	330	3	US-09-640-211A-1290	Seq
27	43.8	5.4	181	3	US-09-513-999C-9489	Seq
28	43.8	5.4	313	3	US-09-640-211A-1350	Seq
29	43.8	5.4	315	3	US-09-640-211A-1348	Seq
30	43.8	5.4	322	3	US-09-640-211A-1337	Seq
31	43.8	5.4	368	3	US-09-640-211A-1379	Seq
32	42.2	5.2	2089	3	US-10-104-047-106	Seq
33	41.6	5.2	259	3	US-09-313-294A-161	Seq
34	38.8	4.8	320	3	US-09-640-211A-1371	Seq
35	38.6	4.8	990	3	US-09-252-991A-1499	Seq
36	38.6	4.8	1131	3	US-09-252-991A-1675	Seq
37	37.8	4.7	290	3	US-09-313-294A-5271	Seq
38	37	4.6	467	3	US-09-640-211A-1667	Seq
39	37	4.6	533	3	US-09-640-211A-1804	Seq
40	36.8	4.6	278	3	US-09-313-294A-1111	Seq
41	36.6	4.6	601	3	US-09-949-016-158118	Seq
42	36.6	4.6	24070	3	US-09-949-016-16153	Seq
43	36.2	4.5	2702	3	US-08-508-761B-3	Seq
44	35.4	4.4	473	3	US-09-640-211A-1740	Seq
45	34.8	4.3	7336	3	US-09-949-016-13935	Seq

ALIGNMENTS

RESULT 1					
US-09-828-303-12					
Sequence 12, Application US/09828303					
Patent No. 6677504					
GENERAL INFORMATION:					
APPLICANT: COSTA E SILVA, OSWALDO DA					
APPLICANT: BOHNER, HANS J.					
APPLICANT: VAN THIELEN, NOCHA					
APPLICANT: CHEN, ROUYING					
TITLE OF INVENTION: TRANSCRIPTION FACTOR STRESS-RELATED PROT					
TITLE OF INVENTION: METHODS OF USE IN PLANTS					
FILE REFERENCE: 16313-0030					
CURRENT APPLICATION NUMBER: US/09/828,303					
CURRENT FILING DATE: 2001-08-20					
PRIOR APPLICATION NUMBER: 60/196,001					
PRIOR FILING DATE: 2000-04-07					
NUMBER OF SEQ ID NOS: 79					
SOFTWARE: Patentin Ver. 2.1					
SEQ ID NO 12					
LENGTH: 804					
TYPE: DNA					
ORGANISM: Physcomitrella patens					
US-09-828-303-12					
Query Match					
Best Local Similarity 100.0%; Pred. No. 1,36-253;					
Matches 804; Conservative 0; Mismatches 0; Indels 0;					
QY	1	ATCCGGGACACGATCCCGCTTAGTGTGTGTCTATTAGTGTGCTG	100.0%;	Score 804;	DB 3; Length 80
DB	1	ATCCGGGACACGATCCCGCTTAGTGTGTGTCTATTAGTGTGCTG			
QY	61	CTTGAGCGAGATTGTCAGAGATTTCTCATPAGCTTCTATAGGAAGA			120
DB	61	CTTGAGCGAGATTGTCAGAGATTTCTCATPAGCTTCTATAGGAAGA			120
QY	121	TAGTGTGTTAAAGATGGCAGCAGCGTGTCTCAGAGAGACCTCG			180
DB	121	TAGTGTGTTAAAGATGGCAGCAGCGTGTCTCAGAGAGACCTCG			180
QY	181	GGGTTCAGTTATGTGCAAGAACTTTGGCGCTTTCTGGCAGCAAGCT			240
DB	181	GGGTTCAGTTATGTGCAAGAACTTTGGCGCTTTCTGGCAGCAAGCT			240
QY	241	GTGCTCGAAGTGCTACCGAGAGACGATCATGACGGAAGTGAAGCT			300
DB	241	GTGCTCGAAGTGCTACCGAGAGACGATCATGACGGAAGTGAAGCT			300

/ OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-7446

Query Match 12.0%; Score 96.8; DB 3; Length 283;
Best Local Similarity 67.2%; Pred. No. 2e-21;
Matches 137; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

QY 509 CCCAATCGATGGGATCTTGACGAGACGCTGTTGATGACGATTTAAGTGTGCTGT 568
DB 2 CCAACCGGTGGCAACCTGTAGGAACGCTGTGGTTACCGGTTTAACTGCGCATGC 61
QY 569 GGCACCTCTACTACGCTTAACTCGCTACTCGACAAACACACTGACATATGACTAC 628
DB 62 GGGAAACACTACTGTTGATGACACCGCTACTCGACAAACACAGACTGCCAGTTGACTAT 121
QY 629 AAAGCCGACGGGACAGAACGATGCGAAAGCTTAATCTTGTGTCGCGCGAAGAGTT 688
DB 122 CGAATCGACGATAGGACGCTATCGCCAAAGCCCAATCCAGTGTGAGAGCGGAGAAAGCTT 181
QY 689 GTCAAGTTTGTATGACATCCGTT 712
DB 182 GACAAGATCTGAGCGGGGCGGCAAT 205

RESULT 4
US-09-313-294A-1951
Sequence 1951, Application US/09313294A

GENERAL INFORMATION:
APPLICANT: Lalqudi, Raghunath V.
APPLICANT: Ito, Laura Y.
APPLICANT: Sherman, Bradley K.
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
FILE REFERENCE: PL-0017 US
CURRENT APPLICATION NUMBER: US/09/313,294A
CURRENT FILING DATE: 1999-05-14
NUMBER OF SEQ ID NOS: 7600
SOFTWARE: PERL Program
SEQ ID NO 1951
LENGTH: 281
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. 6476212 700551796H1
US-09-313-294A-1951

Query Match 11.6%; Score 93.6; DB 3; Length 281;
Best Local Similarity 67.3%; Pred. No. 2.3e-20;
Matches 132; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 509 CCCAATCGATGGGATCTTGACGAGACGCTGTTGATGACGATTTAAGTGTGCTGT 568
DB 10 CCAGACCGGTGGCGCGCTGCGAAGAGCTGTGGCTTACCGGATTTAGTCCGATGC 69
QY 569 GGCACCTCTACTACGCTTAACTCGCTACTCGACAAACACACTGACATATGACTAC 628
DB 70 GGGAACTATGATCTGTTGCTGTCACCGCTACTCGACAAACACAGACTGTGATGACTAT 129
QY 629 AAAGCCGACGGGACAGAACGATGCGAAAGCTTAATCTTGTGTCGCGCGAAGAGTT 688
DB 130 CGGATCTGACGACAGGACCGCATGCGCAAGCCCAATCTCTGTGTGAGGCGGAGAAAGCTC 189
QY 689 GTCAAGTTTGTATGAG 704
DB 190 GACAAGATCTGAGGCG 205

RESULT 5
US-09-313-294A-3856
Sequence 3856, Application US/09313294A
Patent No. 6476212
GENERAL INFORMATION:

APPLICANT: Lalqudi, Raghunath V.

APPLICANT: Ito, Laura Y.

APPLICANT: Sherman, Bradley K.

TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVE

FILE REFERENCE: PL-0017 US

CURRENT APPLICATION NUMBER: US/09/313,294A

CURRENT FILING DATE: 1999-05-14

NUMBER OF SEQ ID NOS: 7600

SOFTWARE: PERL Program

SEQ ID NO 3856

LENGTH: 211

TYPE: DNA

ORGANISM: Zea mays

FEATURE:

NAME/KEY: misc feature

OTHER INFORMATION: Incyte ID No. 6476212 700282265H2

NAME/KEY: unsure

LOCATION: 41, 60, 99, 116, 118, 135, 154

OTHER INFORMATION: a, t, c, g, or other

US-09-313-294A-3856

Query Match 9.8%; Score 78.8; DB 3; Length 21
Best Local Similarity 64.5%; Pred. No. 1.4e-15;
Matches 127; Conservative 0; Mismatches 69; Indels

QY 509 CCCAATCGATGGGATCTTGACGAGACGCTGTTGATGACGATTTA 568
DB 9 CCAGACCGGTGGCGCGCTGCGAAGAGCTGTGGCTGACCGGATTTA 68
QY 569 GGCACCTCTACTACGCTTAACTCGCTACTCGACAAACACACTGTC 627
DB 69 GGGAACTATGATCTGTTGCTGTCACCGCTACTCGACAAACACAGACTAT 128
QY 628 CAAGCCGACGGGACAGAACGATGCGAAAGCTTAATCTTGTGCTGT 687
DB 129 TCGACATCGACGAGAGAGCGGATTTCCAGGCGCAATCTGTGTGAGG 188
QY 688 TGTCAAGTTTGTATGAG 704
DB 189 GACAAGATCTGAGGCG 205

RESULT 6

US-08-861-269-4

Sequence 4, Application US/08861269

Patent No. 5817494

GENERAL INFORMATION:

APPLICANT: Bandman, Olga

APPLICANT: Corley, Neil C.

APPLICANT: Lal, Preeti

TITLE OF INVENTION: UBIQUITIN CONJUGATION PROTEINS

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/861,269

FILING DATE: Filed Herewith

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0302 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1197 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BRAINOT09
CLONE: 2151473
US-08-861-269-4

Query Match 9.6%; Score 77; DB 2; Length 1197;
Best Local Similarity 63.0%; Pred. No. 1.4e-14;
Matches 119; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY 512 AATGATCGGATCTTGACAGAAAGCGTGTGATGACAGATTAACTGCGCTGGC 571
DB 700 AATGCTGTTTCATGTGACAGAAAGTGAGACTTACGCTTGAATGCCGAGTGGA 759
QY 572 AACCTTACTGCGCTTACATCGGTACTCGACAAACACTTGACATATGACTACAA 631
DB 760 AATGTTACTGTGTGTACACCGTTACTCAGATGACACAAATGCTTTACATTACAA 819
QY 632 GCCGACGGCAGAAAGCGATTGCGAAAGCTAATCTTGTGTCGCCGAGAGGTGTC 691
DB 820 GCCGATGCTGTGAGAAATCAGAAAAGAAATCCAGTAGTGTGTGAAAAAGATCCAA 879
QY 692 AAGTTTGA 700
DB 880 AAGATTGA 888

RESULT 7
US-09-134-596-4
Sequence 4, Application US/09134596
Patent No. 592318

GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Corley, Neil C.
APPLICANT: Lal, Preeti
APPLICANT: Shah, Purvi
TITLE OF INVENTION: UBIQUITIN CONJUGATION PROTEINS
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/134,596
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/861,269
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0302 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166

INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1197 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BRAINOT09
CLONE: 2151473
US-09-134-596-4

Query Match 9.6%; Score 77; DB 2; Length 1197;
Best Local Similarity 63.0%; Pred. No. 1.4e-14;
Matches 119; Conservative 0; Mismatches 70; Indels 0;

QY 512 AATGATCGGATCTTGACAGAAAGCGTGTGATGACAGATTAACTG 571
DB 700 AATGCTGTTTCATGTGACAGAAAGTGAGACTTACGCTTGAATG 759
QY 572 AACCTTACTGCGCTTACATCGGTACTCGACAAACACTTGACAT 631
DB 760 AATGTTACTGTGTGTACACCGTTACTCAGATGACACAAATGCT 819
QY 632 GCCGACGGCAGAAAGCGATTGCGAAAGCTAATCTTGTGTCGCC 691
DB 820 GCCGATGCTGTGAGAAATCAGAAAAGAAATCCAGTAGTGTGTG 879
QY 692 AAGTTTGA 700
DB 880 AAGATTGA 888

RESULT 8
US-09-293-273-4
Sequence 4, Application US/09293273
Patent No. 6057112

GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Corley, Neil C.
APPLICANT: Lal, Preeti
APPLICANT: Shah, Purvi
TITLE OF INVENTION: UBIQUITIN CONJUGATION PROTEINS
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/293,273
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/134,596
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0302 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-845-4166
TELEFAX: 415-855-0555
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1197 base pairs
TYPE: nucleic acid
STRANDEDNESS: single

TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BRAINOT09
CLONE: 2151473
US-09-293-273-4

Query Match
Best Local Similarity 9.6%; Score 77; DB 3; Length 1197;
Matches 119; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY 512 ATTCGATGCGATCTTGCAGAGAGCGTGTGATTCAGAGATTAAAGTGTGCTGTGC 571
DB 700 ATTCGCTGTTTATGTGAGAGAAAGTGGAGCTTACGTGGTTGAATGCCGTGTGA 759
QY 572 AACCTTACTGCGCTTAACTGGTACTGCGAACAACACTTGCACATATGACTACAA 631
DB 760 AATGTTACTGTGTGTACACCGTTACAGATGTACAAATGGCTTACATATACAA 819
QY 632 GCGGACGGCAGAGCGATTCGAAAGCTATCTCTGTGTGCGCGAGAGGTGTG 691
DB 820 GCGGATGCTGTGAGAAATTCAGAAAGAAATCCAGTATGTTGTGTAAGATCCAA 879
QY 692 AAGTTTGA 700
DB 880 AAGATTGA 888

RESULT 9
US-09-949-016-810
Sequence 810, Application US/09949016
Patent No. 6812339

GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 810
LENGTH: 2425
TYPE: DNA
ORGANISM: Human
US-09-949-016-810

Query Match
Best Local Similarity 9.2%; Score 74.2; DB 3; Length 2425;
Matches 118; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 498 CTAGACCCGAGCCCATTCGATCGGATCTTGCAGAGAGCGTGTGATTCAGAGATTAA 557
DB 728 CCAACCAAGAAAGAAACAGATGTTTCATGTGAGAAAGAAAGTGTCTTACAGGGTTTG 787
QY 558 AGTGTGCTGTGGCACTCTACTGCGCTTATCTGATCTGCGAACAACACTTGA 617
DB 788 ACTGCCGATGTGAAATTTGTTTGTGACCTTCACCGTTACTGTGCAAGCAACTGTC 847
QY 618 CATATGACTACAAAGCGCAGGCGAGAGCGATTTCGAAAGCTATCTCTGTGTGCG 677
DB 848 CGTATGATTACAAAGAGAGAGCTGCGCAAAATTCAGAAAGAGATCCAGTTGTGTG 907
QY 678 CCGAGAGGTT 688
DB 908 CTGAAAAAATT 918

RESULT 10
US-09-949-016-3693

Sequence 3693, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3693
LENGTH: 2684
TYPE: DNA
ORGANISM: Human
US-09-949-016-3693

Query Match
Best Local Similarity 9.2%; Score 74.2; DB 3; Length 26
Matches 118; Conservative 0; Mismatches 73; Indels 0;

QY 498 CTAGACCCGAGCCCATTCGATCGGATCTTGCAGAGAGCGTGTGATT 557
DB 987 CCAACCAAGAAAGAAACAGATGTTTCATGTGAGAAAGAAAGTGTCT 1046
QY 558 AGTGTGCTGTGGCACTCTACTGCGCTTATCATCTGCTTACCTGAGCA 617
DB 1047 ACTGCCGATGTGAAATTTGTTTGTGACCTTCACCGTTACTGTACAA 1106
QY 618 CATATGACTACAAAGCGCAGGCGAGAGCGATTTCGAAAGCTATTC 677
DB 1107 CGTATGATTACAAAGAGAGAGCTGCGCAAAATTCAGAAAGAGATTC 1166
QY 678 CCGAGAGGTT 688
DB 1167 CTGAAAAAATT 1177

RESULT 11
US-09-244-805-10

Sequence 10, Application US/09244805
Patent No. 6699660
GENERAL INFORMATION:
APPLICANT: Worley, Paul F.
APPLICANT: Lananah, Anthony
APPLICANT: Goetz, Bernard
APPLICANT: Helmsch, Holger
APPLICANT: Kuefer, Rohini
APPLICANT: Scheek, Sigrid
APPLICANT: Nikolich, Karoly
APPLICANT: Zhukovski, Eugene
TITLE OF INVENTION: IMMEDIATE EARLY GENES AND METHODS OF USE;
FILE REFERENCE: 10496/004001
CURRENT APPLICATION NUMBER: US/09/244,805
CURRENT FILING DATE: 1999-02-05
PRIOR APPLICATION NUMBER: 60/074,518
PRIOR FILING DATE: 1998-02-12
PRIOR APPLICATION NUMBER: 60/074,135
PRIOR FILING DATE: 1998-02-06
NUMBER OF SEQ ID NOS: 62
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10
LENGTH: 2515
TYPE: DNA

ORGANISM: Eukaryote
FEATURE:
NAME/KEY: CDS
LOCATION: (414)...(1055)
US-09-244-805-10

Query Match 8.9%; Score 71.8; DB 3; Length 2515;
Best Local Similarity 61.5%; Pred. No. 1e-12;
Matches 115; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 498 CTAGACCCGAGCCCAATGATCGGATCTTTCAGAGAGCGCTTGATGACAGATTGA 557
DB 853 CCAAAACCAAGAAGAACAGATTTTATGCTGTAAGAAAGTTGCGCTTACAGCGTTTG 912
QY 558 AGTGCCTGCTGGCAACCTCTACGCGCTTACATCGGACTCGAGCAACAACACTTGCA 617
DB 913 ACTGCCGAGTGGAAATTTGTTTGTGACCTTCACCGTTACTCTACAGCAACACTGTC 972
QY 618 CATATGACTACAAAGCCGAGGAGGAGCAAGCGATTGCGAAAGCTATCCTCTTGCTGG 677
DB 973 CTATGATTTACAAAGACAGAGAGCTGCAGCAAAAATCAGAAAATCCAGTTGTTGG 1032
QY 678 CCGAGAA 684
DB 1033 CTGAAA 1039

RESULT 12
US-09-949-016-12552
Sequence 12552, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12552
LENGTH: 14524
TYPE: DNA
ORGANISM: Human
US-09-949-016-12552

Query Match 7.6%; Score 61; DB 3; Length 14524;
Best Local Similarity 64.5%; Pred. No. 8.8e-09;
Matches 91; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 548 ACAGATTTAAGTGTGCTGGCAACCTCTACTGCGCTTACATCGGTAATCGAGCAAA 607
DB 10877 ACAGGTTTGAATGCGGAGTGGAAATTTGTTTGTGACCTTACTACTACAG 10936
QY 608 CACACTTGACATGATGACTCAAAAGCCGAGGAGGAGAAAGCATGGCGAAAGCTAATCCT 667
DB 10937 CACAACTGTCCGTATGATTTACAAAGCAGAGAGAGCTGCAGCAAAAATCAGAAAGAGATCCA 10996
QY 668 CTGTGCTGGCCGAGAGGTT 688
DB 10997 GTTGTGTGGCTGAAAAAATT 11017

RESULT 13
US-09-949-016-15435
Sequence 15435, Application US/09949016
Patent No. 6812339

GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 15435
LENGTH: 14785
TYPE: DNA
ORGANISM: Human
US-09-949-016-15435

Query Match 7.6%; Score 61; DB 3; Length 1478
Best Local Similarity 64.5%; Pred. No. 8.8e-09;
Matches 91; Conservative 0; Mismatches 50; Indels 0;

QY 548 ACAGATTTAAGTGTGCTGGCAACCTCTACTGCGCTTACATCGG 607
DB 11137 ACAGGTTTGAATGCGGAGTGGAAATTTGTTTGTGACTTCACTGCTT 11196
QY 608 CACACTTGACATATGACTCAAAAGCCGAGGAGGAGCAAGCATTTGCA 667
DB 11197 CACAACTGTCCGTATGATTTACAAAGCAGAGAGCTGCAGCAAAAATCAGAA 11256
QY 668 CTGTGCTGGCCGAGAGGTT 688
DB 11257 GTTGTGTGGCTGAAAAAATT 11277

RESULT 14
US-09-640-211A-1742
Sequence 1742, Application US/09640211A
Patent No. 6833446
GENERAL INFORMATION:
APPLICANT: Wood, Marion
APPLICANT: Shenk, Michael A.
APPLICANT: McGrath, Annette
APPLICANT: Glenn, Matthew
TITLE OF INVENTION: Compositions and Methods for the
TITLE OF INVENTION: Modification of Gene Transcription
FILE REFERENCE: 11000.1021CU
CURRENT APPLICATION NUMBER: US/09/640,211A
CURRENT FILING DATE: 2000-08-16
NUMBER OF SEQ ID NOS: 2368
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1742
LENGTH: 348
TYPE: DNA
ORGANISM: Pinus radiata
US-09-640-211A-1742

Query Match 7.4%; Score 59.4; DB 3; Length 34
Best Local Similarity 63.8%; Pred. No. 4.4e-09;
Matches 90; Conservative 0; Mismatches 51; Indels 0;

QY 512 AATGATCGGATCTTGCAGAGAGCGTTGATTTGACAGGATTTAAGT 571
DB 204 AAGAGGTGTTTGTGCTGTAACAGCGGGGTGGGCTGCAGCGGCTTTAAGT 263
QY 572 AACCTTACTGCGCTTTACATCGGTAATCGGACAAACACACTTGACAT 631
DB 264 GACCTTTTCTGCGCTCAGCAGAGGTAATCTGATATGATGACTGCTCTT 323
QY 632 GCGCAGGAGGAGAGGAT 652

Db 324 ACTGCCGCGCCTCGCAT 344

RESULT 15
US-09-313-294A-4836
; Sequence 4836, Application US/09313294A

; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Laljudi, Raghunath V.

; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR

; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14

; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 4836
; LENGTH: 277

; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:

; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6476212 700349114H1
; NAME/KEY: unsure
; LOCATION: 18, 21

; OTHER INFORMATION: a, t, c, g, or other
; US-09-313-294A-4836

Query Match
Best Local Similarity 61.7%; Score 54.6; DB 3; Length 277;
Matches 87; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 156 AGGAGCGACTCGCAGGCGCTGAGGCTCACTTATGTGCAAGAACCTTGGCGCTCT 215
DB 122 AGGAGCGCTGGTGCAGCGCCCGAGGAGCCATCCTCTGCAATCAATACTGTGGCTTCT 181
QY 216 TGGCAGCCCAAGCTACATGAGGCTGTGCTCGAAGTCTACCGAGACAGTCATGCAAG 275
DB 182 TGGCAGCGCGCGGCGACATGATGCTCCAGTGCACCAAGAGATGATTAACGAGC 241
QY 276 CGAAGTACGCGCTTTAGCTG 296
DB 242 AGGATCAGGCCAAGCTGCTG 262

Search completed: December 8, 2005, 10:15:34
Job time : 200 secs

This Page Blank (uspto)

Db 241 GTGCTGAAGTGTCTACCGAGAGACAGTCATCGCAAGCAAGAGTGAAGCTTTAGCTGACGA 300
Qy 301 AGCCACTAGAGGTGTCTCAGGCGGACATCTGCGCAACACTGTGTCTGTTACGCCCCCGCTCC 360
Db 301 AGCCACTAGAGGTGTCTCAGGCGGACATCTGCGCAACACTGTGTCTGTTACGCCCCCGCTCC 360
Qy 361 TGTACATGAGACCAAGCTTCACATGCGAGGTTGAGAAACAATGATTTGTGCCGATCAATC 420
Db 361 TGTACATGAGACCAAGCTTCACATGCGAGGTTGAGAAACAATGATTTGTGCCGATCAATC 420
Qy 421 TTCCAGCTATCAACAAGACCTGTGTTACCCCGCTGCAAGCTGCTCCCTCAGGCAGTGAAGTC 480
Db 421 TTCCAGCTATCAACAAGACCTGTGTTACCCCGCTGCAAGCTGCTCCCTCAGGCAGTGAAGTC 480
Qy 481 CTCTATGCGAGTCTCCTCTAGACCCGAGCCCAATGAGCGGATCTTGAGAGAAACGAGT 540
Db 481 CTCTATGCGAGTCTCCTCTAGACCCGAGCCCAATGAGCGGATCTTGAGAGAAACGAGT 540
Qy 541 TGGATTGACAGGATTTAAAGTGTGCTGCGCAACCTCTACTGCGCTTTACATCGGTACTC 600
Db 541 TGGATTGACAGGATTTAAAGTGTGCTGCGCAACCTCTACTGCGCTTTACATCGGTACTC 600
Qy 601 GGACAAACAACACTTGCACATATGACTACAAAGCCGAGGGCAGAAAGCGATTGCGAAAGC 660
Db 601 GGACAAACAACACTTGCACATATGACTACAAAGCCGAGGGCAGAAAGCGATTGCGAAAGC 660
Qy 661 TTAATCTCTGTGCGGCGCGAGAAAGTTGTCAAGTTTGTGATGAGATCCGTTAAACCTTTT 720
Db 661 TTAATCTCTGTGCGGCGCGAGAAAGTTGTCAAGTTTGTGATGAGATCCGTTAAACCTTTT 720
Qy 721 CTGCCGACGATTTAGGCTTCATACATTTGAGTAACCTTCTTCTTTATCGAAG 780
Db 721 CTGCCGACGATTTAGGCTTCATACATTTGAGTAACCTTCTTCTTTATCGAAG 780
Qy 781 AGCGAGTGCATCAAGAGCTCGCC 804
Db 781 AGCGAGTGCATCAAGAGCTCGCC 804

RESULT 2
US-10-716-089-12
; Sequence 12, Application US/10716089
; Publication No. US20040107463A1
; GENERAL INFORMATION:
; APPLICANT: COSTA E SILVA, OSWALDO DA
; APPLICANT: BOHNER, HANS J.
; APPLICANT: VAN THIELEN, NOCHA
; APPLICANT: CHEN, ROUYING
; TITLE OF INVENTION: TRANSCRIPTION FACTOR STRESS-RELATED PROTEINS AND
; FILE REFERENCE: 16313-0030
; CURRENT APPLICATION NUMBER: US/10/716,089
; CURRENT FILING DATE: 2003-11-18
; PRIOR APPLICATION NUMBER: 60/196,001
; PRIOR FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 12
; LENGTH: 804
; TYPE: DNA
; ORGANISM: Physcomitrella patens

Query Match 100.0%; Score 804; DB 7; Length 804;
Best Local Similarity 100.0%; Pred. No. 1,2e-265; Indels 0; Gaps 0;
Matches 804; Conservative 0; Mismatches 0;

Qy 1 ATCCGGGACACGATCCGCTTAGTGTGTGTCATTAAGTTGTTGTTGCAAGTGAAGC 60
Db 1 ATCCGGGACACGATCCGCTTAGTGTGTGTCATTAAGTTGTTGTTGCAAGTGAAGC 60
Qy 61 CTGAGCGAGATTTGCAAGATTTTCTCATACGCTTCTGATTAAGAAAGTACACCTTAT 120
Db 61 CTGAGCGAGATTTGCAAGATTTTCTCATACGCTTCTGATTAAGAAAGTACACCTTAT 120

Db 61 CTGAGCGAGATTTGCAAGATTTTCTCATACGCTTCTGATTAAGAAAGT 120
Qy 121 TAGTCTGTTAAAGATGGCCACCGAGCGGTGTCTCAGGAGCGACTGTG 180
Db 121 TAGTCTGTTAAAGATGGCCACCGAGCGGTGTCTCAGGAGCGACTGTG 180
Qy 181 GGGTTCAGTTATGTGCAAGAACTTTGCGGCTTCTTCCGACCGAAGCT 240
Db 181 GGGTTCAGTTATGTGCAAGAACTTTGCGGCTTCTTCCGACCGAAGCT 240
Qy 241 GTGCTGGAAGTGTCTACCGAGAGACATGTATGCAAGCGAAGTGAAGCT 300
Db 241 GTGCTGGAAGTGTCTACCGAGAGACATGTATGCAAGCGAAGTGAAGCT 300
Qy 301 AGCCACTAGAGGTGTCTCAGGCGGACATCTGCGCAACACTGTGTCTGTTACGCCCCCGCTCC 360
Db 301 AGCCACTAGAGGTGTCTCAGGCGGACATCTGCGCAACACTGTGTCTGTTACGCCCCCGCTCC 360
Qy 361 TGTACATGAGACCAAGCTTCACATGCGAGGTTGAGAAACAATGATTTGTG 420
Db 361 TGTACATGAGACCAAGCTTCACATGCGAGGTTGAGAAACAATGATTTGTG 420
Qy 421 TTCCAGCTATCAACAAGACCTGTGTTACCCCGCTGCAAGCTGCTCCCTCAG 480
Db 421 TTCCAGCTATCAACAAGACCTGTGTTACCCCGCTGCAAGCTGCTCCCTCAG 480
Qy 481 CTCTATGCGAGTCTCCTCTAGACCCGAGCCCAATGAGCGGATCTTGAGAGAAACGAGT 540
Db 481 CTCTATGCGAGTCTCCTCTAGACCCGAGCCCAATGAGCGGATCTTGAGAGAAACGAGT 540
Qy 541 TGGATTGACAGGATTTAAAGTGTGCTGCGCAACCTCTACTGCGCTTTA 600
Db 541 TGGATTGACAGGATTTAAAGTGTGCTGCGCAACCTCTACTGCGCTTTA 600
Qy 601 GGACAAACAACACTTGCACATATGACTACAAAGCCGAGGGCAGAAAGCG 660
Db 601 GGACAAACAACACTTGCACATATGACTACAAAGCCGAGGGCAGAAAGCG 660
Qy 661 TTAATCTCTGTGCGGCGCGAGAAAGTTGTCAAGTTTGTGATGAGCATCC 720
Db 661 TTAATCTCTGTGCGGCGCGAGAAAGTTGTCAAGTTTGTGATGAGCATCC 720
Qy 721 CTGCCGACGATTTAGGCTTCATACATTTGAGTAACCTTCTTCTTTATCGAAG 780
Db 721 CTGCCGACGATTTAGGCTTCATACATTTGAGTAACCTTCTTCTTTATCGAAG 780
Qy 781 AGCGAGTGCATCAAGAGCTCGCC 804
Db 781 AGCGAGTGCATCAAGAGCTCGCC 804

RESULT 3
US-09-828-303-4
; Sequence 4, Application US/09828303
; Patent No. US20020102695A1
; GENERAL INFORMATION:
; APPLICANT: COSTA E SILVA, OSWALDO DA
; APPLICANT: BOHNER, HANS J.
; APPLICANT: VAN THIELEN, NOCHA
; APPLICANT: CHEN, ROUYING
; TITLE OF INVENTION: TRANSCRIPTION FACTOR STRESS-RELATED PROT
; FILE REFERENCE: 16313-0030
; CURRENT APPLICATION NUMBER: US/09/828,303
; CURRENT FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: 60/196,001
; PRIOR FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 4
; LENGTH: 1531
; TYPE: DNA
; ORGANISM: Physcomitrella patens

US-09-828-303-4

Query Match 96.6%; Score 776.4; DB 3; Length 1531;
Best Local Similarity 99.7%; Pred. No. 5,4e-256;
Matches 788; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 8 GCACCAAGTCCCGCTAGTGTGTGTGTCATTTAGTGTGTGCAAGCTGCAAGCTTGAGC 67
DB 529 GCACCAAGTCCCGCTAGTGTGTGTGTCATTTAGTGTGTGCAAGCTGCAAGCTTGAGC 588
DB 68 GAGATTGCAAGATTCTTCATACGCTTCTGATTAGAAAAGATACCCCTTATTAGTCTG 127
DB 589 GAGATTGCAAGATTCTTCATACGCTTCTGATTAGAAAAGATACCTTATTAGTCTG 648
QY 128 TTAAGATGGCCAGCGAGGTGTGTCTAGAGAGACCACTCCCAAGGCCCTTGAAGGTCA 187
DB 649 TTAAGATGGCCAGCGAGGTGTGTCTAGAGAGACCACTCCCAAGGCCCTTGAAGGTCA 708
QY 188 GTTATGTGCAGAACCTTTGCGGCTTCTTCCGACGCAAGCTACCATGAGGTTGTGCTCG 247
DB 709 GTTATGTGCAGAACCTTTGCGGCTTCTTCCGACGCAAGCTACCATGAGGTTGTGCTCG 768
QY 248 AAGTCTACCGAGAGACAGTCAATGCAAGCGAAGATGACGGCTTTAGCTGAGCAAGCACT 307
DB 769 AAGTCTACCGAGAGACAGTCAATGCAAGCGAAGATGACGGCTTTAGCTGAGCAAGCACT 827
QY 308 CAGGCTGCTCAGGCGACATCTGCCACAGCTGCTGCTTCCACCCCGCTCTGTACT 367
DB 828 CAGGCTGCTCAGGCGACATCTGCCACAGCTGCTGCTTCCACCCCGCTCTGTACT 887
QY 368 GAGACCAAGCTCAGTCGCGAGGTGAGAAACATGATGTGCCGATCAATCTTCCAGC 427
DB 888 GAGACCAAGCTCAGTCGCGAGGTGAGAAACATGATGTGCCGATCAATCTTCCAGC 947
QY 428 TATCAACAAGACTGTGTTACCCCGCTGCAAGCTGCTCCCTCAGGCAAGTAACTCTATC 487
DB 948 TATCAACAAGACTGTGTTACCCCGCTGCAAGCTGCTCCCTCAGGCAAGTAACTCTATC 1007
QY 488 GCAGTCCCTCTAGACCCGAGCCCAATGATGCGGATCTTGAAGAAAGCTGTGATG 547
DB 1008 GCAGTCCCTCTAGACCCGAGCCCAATGATGCGGATCTTGAAGAAAGCTGTGATG 1067
QY 548 ACAGGATTTAAGTGTGCTGTGGCAACCTCTACTGCTTTTACATCGGATCTGAGCAAA 607
DB 1068 ACAGGATTTAAGTGTGCTGTGGCAACCTCTACTGCTTTTACATCGGATCTGAGCAAA 1127
QY 608 CACACTTGACATATGACTACAAAGCCGAGGGCAGAAAGCGATTGCCAAAGCTAATCT 667
DB 1128 CACACTTGACATATGACTACAAAGCCGAGGGCAGAAAGCGATTGCCAAAGCTAATCT 1187
QY 668 CTGTGCGTGGCCGAGAAAGTTGTCAAGTTTGTATGAGCAATCGGTTAAGCTTTTCTGCCGA 727
DB 1188 CTGTGCGTGGCCGAGAAAGTTGTCAAGTTTGTATGAGCAATCGGTTAAGCTTTTCTGCCGA 1247
QY 728 CGATTAGGCTTCAATATGAGTAACTCTACATCTTCTTATGAGAGAGCGAGT 787
DB 1248 CGATTAGGCTTCAATATGAGTAACTCTACATCTTCTTATGAGAGAGCGAGT 1307
QY 788 CGCATCAAGA 797
DB 1308 CGCATCAAGA 1317

RESULT 4

US-10-716-089-4
; Sequence 4, Application US/10716089
; Publication No. US20040107463A1
; GENERAL INFORMATION:
; APPLICANT: COSTA E SILVA, OSWALDO DA
; APPLICANT: BOHNER, HANS J.
; APPLICANT: VAN THIELEN, NOCHA
; APPLICANT: CHEN, ROUING
; TITLE OF INVENTION: TRANSCRIPTION FACTOR STRESS-RELATED PROTEINS AND

TITLE OF INVENTION: METHODS OF USE IN PLANTS

FILE REFERENCE: 16313-0030
CURRENT APPLICATION NUMBER: US/10/716,089
CURRENT FILING DATE: 2003-11-18
PRIOR APPLICATION NUMBER: 60/196,001
PRIOR FILING DATE: 2000-04-07
NUMBER OF SEQ ID NOS: 79
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 4
LENGTH: 1531
TYPE: DNA
ORGANISM: Physcomitrella patens
US-10-716-089-4

Query Match 96.6%; Score 776.4; DB 7; Length 1
Best Local Similarity 99.7%; Pred. No. 5,4e-256;
Matches 788; Conservative 0; Mismatches 1; Indels 1;

QY 8 GCACCAAGTCCCGCTAGTGTGTGTGTCATTTAGTGTGTGCAAGCTGCAAGCTTGAGC 67
DB 529 GCACCAAGTCCCGCTAGTGTGTGTGTCATTTAGTGTGTGCAAGCTGCAAGCTTGAGC 588
QY 68 GAGATTGCAAGATTCTTCATACGCTTCTGATTAGAAAAGATACACCC 127
DB 589 GAGATTGCAAGATTCTTCATACGCTTCTGATTAGAAAAGATACATCC 648
QY 128 TTAAGATGGCCAGCGAGGTGTGTCTAGAGAGACCACTCCGAGGCC 187
DB 649 TTAAGATGGCCAGCGAGGTGTGTCTAGAGAGACCACTCCGAGGCC 708
QY 188 GTTATGTGCAGAACCTTTGCGGCTTCTTCCGACGCAAGCTACCATG 247
DB 709 GTTATGTGCAGAACCTTTGCGGCTTCTTCCGACGCAAGCTACCATG 768
QY 248 AAGTCTACCGAGAGACAGTCAATGCAAGCGAAGATGACGGCTTTAGCTG 307
DB 769 AAGTCTACCGAGAGACAGTCAATGCAAGCGAAGATGACGGCTTTAGCTG 827
QY 308 CAGGCTGCTCAGGCGACATCTGCCACAGCTGCTGCTTCCACCCCGC 367
DB 828 CAGGCTGCTCAGGCGACATCTGCCACAGCTGCTGCTTCCACCCCGC 887
QY 368 GAGACCAAGCTCAGTCGCGAGGTGAGAAACATGATGTGCCGATCA 427
DB 888 GAGACCAAGCTCAGTCGCGAGGTGAGAAACATGATGTGCCGATCA 947
QY 428 TATCAACAAGACTGTGTTACCCCGCTGCAAGCTGCTCCCTCAGGCAAGTAACTCTATC 487
DB 948 TATCAACAAGACTGTGTTACCCCGCTGCAAGCTGCTCCCTCAGGCAAGTAACTCTATC 1007
QY 488 GCAGTCCCTCTAGACCCGAGCCCAATGATGCGGATCTTGAAGAAAGCTGTGATG 547
DB 1008 GCAGTCCCTCTAGACCCGAGCCCAATGATGCGGATCTTGAAGAAAGCTGTGATG 1067
QY 548 ACAGGATTTAAGTGTGCTGTGGCAACCTCTACTGCTTTTACATCGGATCTGAGCAAA 607
DB 1068 ACAGGATTTAAGTGTGCTGTGGCAACCTCTACTGCTTTTACATCGGATCTGAGCAAA 1127
QY 608 CACACTTGACATATGACTACAAAGCCGAGGGCAGAAAGCGATTGCCAAAGCTAATCT 667
DB 1128 CACACTTGACATATGACTACAAAGCCGAGGGCAGAAAGCGATTGCCAAAGCTAATCT 1187
QY 668 CTGTGCGTGGCCGAGAAAGTTGTCAAGTTTGTATGAGCAATCGGTTAAGCTTTTCTGCCGA 727
DB 1188 CTGTGCGTGGCCGAGAAAGTTGTCAAGTTTGTATGAGCAATCGGTTAAGCTTTTCTGCCGA 1247
QY 728 CGATTAGGCTTCAATATGAGTAACTCTACATCTTCTTATGAGAGAGCGAGT 787
DB 1248 CGATTAGGCTTCAATATGAGTAACTCTACATCTTCTTATGAGAGAGCGAGT 1307
QY 788 CGCATCAAGA 797
DB 1308 CGCATCAAGA 1317

RESULT 5
US-10-424-599-99368
; Sequence 99368, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 99368
; LENGTH: 1385
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_60743C.1
US-10-424-599-99368
Query Match 13.3%; Score 106.6; DB 7; Length 1385;
Best Local Similarity 68.2%; Pred. No. 3.8e-25;
Matches 148; Conservative 0; Mismatches 69; Indels 0; Gaps 0;
QY 509 CCCAATCGATGCGATCTTTCAGAGAGCGGTGTGGATTGACAGATTAAAGTGTGCTGT 568
DB 805 CCAATTCATGAGCACTGCAATTAAGCGGTTGGTTGACAGATTAAATTGTGATGT 864
QY 569 GGCAACCTCTACTGCGCTTTTACATGCTGTACTGGCAACAACACACTTGCACATATGACTAC 628
DB 865 GGTAACTTTCTGTTCTGTATCATGCTACTCAGCAAGCATTAATTGCGCATTTGATTAC 924
QY 629 AAAGCGGAGGCGACAGAGCGATTGCGAAAGCTATCCCTTGTGTCGCGGAGAGGT 688
DB 925 CACACTGCTGCAAGAGATGCCATAGCTTAAAGCAATTCGCGCTGTACAGTTGAGAACTT 984
QY 689 GTCAAGTTTGTATGAGCATCCGTTAAGCTTTTCTGCC 725
DB 985 GATTAAGATCTAATTTGTTGCTTTAAGTTTCATGCC 1021
RESULT 6
US-10-425-114-22561
; Sequence 22561, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 22561
; LENGTH: 1003
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3587-277-F2_FLI
US-10-425-114-22561
Query Match 12.9%; Score 103.4; DB 7; Length 1003;
Best Local Similarity 53.6%; Pred. No. 4.1e-24;

Matches 302; Conservative 0; Mismatches 231; Indels 3;
QY 156 AGGAGAGACTCTCCAGAGCCCTTGAGGCTCCAGTTATGTGCAAGACTT 215
DB 122 AGGAGGCTGGGTGCCAGGCCCCGAGGAGACCATTCCTTCATCAATATAC 181
QY 216 TCGGACGCAAGCTACATGAGGGGTTGTGCTGGAAGTGTCAACGAGAGAT 275
DB 182 TCGGACGCGCGGCAATGATGATGATGATGATGATGATGATGATGATGAT 241
QY 276 CGAAGATGACGCGCTTTAGCTGAGCAAGCCACTCAGGCTGCTCAGGCACT 335
DB 242 AGGATCAGGCCCAAGCTGCTGC-----CTCCTCTATGAC 287
QY 336 CTGCTGCTGTTCAAGCCCGCTCCTGATGATGAGACCAACTCAGAG 395
DB 288 CGGACGACGCGCT-----CATGAGCGGTTGTTGCTG 333
QY 396 GAACATGATTTGTCGCGCATCATCTTCACGTTATCAACAGACTGT 455
DB 334 GTAGTAGCTGTGGCCCAAGTTGAGTTGCAACATGAAC--GTGACAGA 391
QY 456 CAGTCCCTTCAGGCACTGAAGTCTCTATGCAAGCTCCCTTAGACC 515
DB 392 TTGCCCGAACCCAGCGAGGGGTGGCGGATCTCCAAAGGGGAGAGGT 451
QY 516 GATCGGATCTTTCAGAGAGCGGTGGATTGACAGATTAAAGTGTG 575
DB 452 GGTGACGCGCTTCGAGAGAGGAGGTGACTTACGGGATTCATGCTGCG 511
QY 576 TCTACTGCGCTTTACATGCTGACTCGACAAACACACTTGCACATATGA 635
DB 512 TGTATGCGACTCCACGCTACTCCGACAGACAGATGCAAGTTGCA 571
QY 636 CAGGCGAGAGAGCATTCGCAAGCTATCTCTTGTGTGTCGCGAGAA 695
DB 572 CTGCCAGGAGCGCATTCGCAAGCTATCTGTTGTGTAAGGACAGACA 631
QY 696 TTTGATGAGCATTCGTTAAGCTT 718
DB 632 TCTAGGCGGGGTCTCTTAGCGGT 654
RESULT 7
US-10-425-114-27861
; Sequence 27861, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Impro
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 27861
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4740-074-B6_FLI
US-10-425-114-27861
Query Match 12.8%; Score 103; DB 7; Length 873
Best Local Similarity 53.2%; Pred. No. 5.2e-24;
Matches 307; Conservative 0; Mismatches 240; Indels 3;
QY 156 AGGAGACGACTCGAGGCGCCCTGAGGCTCCAGTTATGTGCAAGACTT 215

```

Db      144 AGGAGGCTGGGTCCAGGCCCCCGAGGAGCCATCTCTGCAATCAATACTGGGCTTCT 203
Qy      216 TGGGACCCAAAGCTACATGAGGTTGTCTGCAAGTCTACCGAGAGACATGCAAG 275
Db      204 TCGGACGCGCGGCGACATGAAATGTGCTTAAGTCCCAAGAGAGATGATTAAGAAAC 263
Qy      276 CGAAGATGACGGCTTAAAGTGAAGCAACCATGAGCTCTGAGGGAACATCTGCCACAG 335
Db      264 AGGATCAGGCTCAAGCTGGCTGC-----CTCCTTAATGACAGATCGTGA 309
Qy      336 CTGCTGCTGTTCAGCCCCCGCTCTCTGTATCATGAGACCAAGCTACATGAGTTGAGA 395
Db      310 CGGACGCGAGCGCT-----CATGAGCCGGTTGTGTCTGACGCAACACG 355
Qy      336 GAAACAATGATTTGTGCCGATCAATCTTCAAGCTATCAACAAGACCTGTGTTACCCGCTG 455
Db      356 GTATGATAGCTGTGTCCCAAGTCCAGTTGCAAAACATGAAAC--GTGCAGACCGCGCTGATG 413
Qy      456 CAGCTGCCCTCAGGACGTAAGTCTCTATGCGAGCTCCCTCTAGACCCGAGCCCAATC 515
Db      414 TTGCGGACCCAGCGAGGGGGTGGCGGCGATCTCCAAAGGGGGAAGTAAAGGCGGAAAC 473
Qy      516 GATGCGGATCTTTGACGAAGCGTGTGATTTGACAGATTTTAAGTGTGCTGTGCAAC 575
Db      474 GGTGACGCGCTTCGACGAAGAGGTTGGCTCAGCGGATTTCAATGCTCCGTTGGAACT 533
Qy      576 TCTACTGCGCTTTACATCGGTACTCGACAAACACACTTGCACATATGACTACAAAGCCG 635
Db      534 TGTACTGCGCGCTTCACGCGTACTCCGACAAACGACTGCAAGTTGCACTACCGGACTG 593
Qy      636 CAGGCGACGAAGCGATTTGGAAAGCTAATCTTGTGTGCTGCGGAGAGAGTTGTCACT 695
Db      594 CTGCCAGGAGCGCATTTGCAAGGCTAATCCGTTGGTGAAGGCAAGCAAGCTGACAAAGA 653
Qy      696 TTTGATGAGCATCCGTTAACTTTCTGCCGACGAT 732
Db      654 TCTAGGGGGGGTTCCTACGCTTGTCTCAGAAAGATT 690

```

```

RESULT 8
US-10-425-114-26146
; Sequence 26146, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE OF INVENTION: Plants and Uses thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 26146
; LENGTH: 945
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LTB4073-024-B8_FLI
US-10-425-114-26146

```

```

Query Match      12.8%; Score 103; DB 7; Length 945;
Best Local Similarity 52.3%; Pred. No. 5.4e-24;
Matches 284; Conservative 0; Mismatches 220; Indels 39; Gaps 1;
Qy      156 AGAAGACGACCTCGAGGCGCCCTGAAGGTCCAGTTATGTGCAAGAACCTTTGCGGCTTCT 215
Db      113 AGGAGACTGAGATGCGAGGACCTGAGGAGCCATCTTGTGATCAATAATGCGGCTTCT 172

```

```

Qy      216 TGGGACCCAAAGCTACATGAGGTTGTCTGCAAGTCTACCGAGAGAC 275
Db      173 TCGGACGCGGAGCTACCAATGAACATGTGCTCCAAAGTCCCAAGAGAGA 232
Qy      276 CGAAGATGACGGCTTAAAGTGAAGCAACCATGAGCTCTGAGGGAAC 335
Db      233 AGGACGAGCCAAAGCTGGCTG----- 253
Qy      336 CTGCTGCTGTTCAGCCCCCGCTCTCTGTATCATGAGACCAAGCTCAATC 395
Db      254 CCTCTCTATGACAGCATGCTCAATGAGCAAGATGCTGTATGAAAC 313
Qy      396 GAAACAATGATTTGTGCCGATCAATCTTCCAGTATCAACAAGACCTGTG 455
Db      314 GCAACACAGTGTGCTGTGCTCTCAATGAGTGAAGTCAACAATGAAAGT 373
Qy      456 CAGTGCCCTCAGGACGTAAGTCTCTATGCAAGCTCCCTCTAGACC 515
Db      374 ATGTTGCTGGAACCTAGCGAGGGGGCGCGGTATCTCCAAAGGAAAGT 433
Qy      516 GATGCGGATCTTTGACGAAGCGTGTGATTTGACAGATTTTAAGTGTG 575
Db      434 GGTGACGACCTTTCAGGAAGAGGTTGGACTTACAGATTTGAACTGCCG 493
Qy      576 TCTACTGCGCTTTACATGAGTACTCGGACAAACACACTTGCACATATGA 635
Db      494 TGTACTGTGCACTGACCGCTACTCCGACAAACGACGCTGCAAGTTGCA 553
Qy      636 CAGGCGACGAAGCGATTTGGAAAGCTAATCTTGTGTGCTGCGGAGAG 695
Db      554 CTGCTAGGAGATGCCATTTGCCAAAGCTAATCCAGTGTGAAGCGGACAA 613
Qy      696 TTT 698
Db      614 TCT 616

```

```

RESULT 9
US-10-425-115-71431
; Sequence 71431, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules
; FILE REFERENCE: 38-21(53222)B
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 71431
; LENGTH: 947
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MWT4577_16514C.1
US-10-425-115-71431

```

```

Query Match      12.8%; Score 103; DB 8; Length 947
Best Local Similarity 52.3%; Pred. No. 5.4e-24;
Matches 284; Conservative 0; Mismatches 220; Indels 1;
Qy      156 AGAAGACGACCTCGAGGCGCCCTGAAGGTCCAGTTATGTGCAAGAACCT 215
Db      115 AGGAGACTGAGATGCGAGGACCTGAGGAGCCATCTTGTGATCAATTA 174
Qy      216 TCGGACGCGGAGCTACATGAGGTTGTGCTGCAAGTGTCTACCGAGAGAC 275
Db      175 TCGGACGCGGAGCTACATGAAATGTGCTCAAGTGTGCAACAAGAGAT 234
Qy      276 CGAAGATGACGGCTTTAGCTGACGAAGCAAGCACTCAGGCTGCTCAGGCGAC 335

```

```
Db      235 AGGAGCAGCCCAAGTGGCTG----- 255
Qy      336 CTGCTGCTGTTTCAGCCCCCGCTCCTGTACATGAGACCAAGCTCACATGCCAGTTGAGA 395
Db      256 CCTCTCTATGACACAGCATGTCATGTCACAGATGTCATGAGCAACAGTTGTTGCTG 315
Qy      396 GAACAATGATTGCGCGCATCAATTTCCAGCTATCAACAAGACTGGTTACCCCGCTG 455
Db      316 GCACACAGTGTGCTGCTGCTGCTCAATGAGTTGCAACAATGAACGTGACGCCGCTG 375
Qy      456 CAGTCCCTCAGCAGCAGTGAATGCTCTATGCACTCCCTGTAGACCCGACCCCAATC 515
Db      376 ATGTTGCTGAGACTGACGAGGGGGGGGGGTATCTCCAAAGGAAAGTAGGCCGAAAC 435
Qy      516 GATCGGGAATCTTGACAGAAAGCTGTGGATTGACAGATTAAAGTGTGCTGGCAAC 575
Db      436 GGTGACGACCTTGACAGAAAGGTTGGACTTACAGATTCAACTGCCGCTGTGGAACT 495
Qy      576 TCTACTGGGCTTTACATGGGTACTCGGACAAACACACTTGCATATGACTACAAAGCG 635
Db      496 TGTACTGTGACTGACCGCTACTCCGACACAGACTGCAAGTTCCATATCGGACTG 555
Qy      636 CAGGCGAGGAAGCGATTGGGAAGCTAACTCTTGTGCTGGCCGAGAAAGTTGCAAGT 695
Db      556 CTGCTAGGGATGCCATTGGCAGAGCTAATCCAGTGTGAAAGCGGACAAAGCTCGACAA 615
Qy      696 TTT 698
Db      616 TCT 618
```

RESULT 10

```
US-10-767-701-15574
; Sequence 15574, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(5353)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 15574
; LENGTH: 1142
; TYPE: DNA
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS3087_1
US-10-767-701-15574
```

Query Match 12.8%; Score 103; DB 7; Length 1142;

Best Local Similarity 52.3%; Pred. No. 66-24;

Matches 284; Conservative 0; Mismatches 220; Indels 39; Gaps 1;

```
Qy      156 AGGAGCAGCCTCGCAGGCCCTGAGGGTCCAGTTATGTGCAAGAACTTTGCGGCTTCT 215
Db      173 AGGAGACTGATGACGAGCACTGAGGGACCATCTTTGCATCAATAACTCGGCTTCT 232
Qy      216 TGGGAGCCAACTTACATGGGGTGTGTGCAAGTGTCAACGAGACAGTCATGCAAG 275
Db      233 TCGGAGCGCAGCTACATGAAATGTGTCCAAAGTCCCAAGAGATGATTAAGAAC 292
Qy      276 CGAAGATGACGGCTTATGCTGAGCAAGCACTCAGGCTGCTCAGCGGACATCTGCCAC 335
Db      293 AGGAGCAGCCCAAGCTGGCTG----- 313
Qy      336 CTGCTGCTGTTTCAGCCCCCGCTCCTGTACATGAGACCAAGCTCACATGCCAGTTGAGA 395
Db      314 CCTCTCTATGACACAGCATGTCATGAGCAAGATGCTGTCAATGAAACAGTTGTTGCTG 373
```

```
Qy      396 GAACAATGATTGCGCGCATCAATCTTCAGCTATCAACAAGACTGGT 455
Db      374 GCACACAGTGTGCTGCTGCTGCTCAATCGAGTTGCAACAATGAGCT 433
Qy      456 CAGTCCCTCAGCAGTGAATGCTCTATGCACTCCCTTAGACC 515
Db      434 ATGTTGCTGACCTTAGGAGGGGGGGGTGATCTCCAAAGGAAAGT 493
Qy      516 GATCGGGAATCTTGACAGAAAGCTGTGGATTGACAGATTAAGTGTG 575
Db      494 GGTGACGACCTTGACAGAAAGGTTGGACTTACAGGATTCAGCTGCC 553
Qy      576 TCTACTGGGCTTTACATGGAATCGGACTCGGACAAACACACTTGCATATGA 635
Db      554 TGTACTGTGACTGACCGCTACTCCGACACAGACTGCAAGTTGCA 613
Qy      636 CAGGCGAGGAAGCGATTGGCAAGCTAATCTCTTGTGTCGGCGAGAA 695
Db      614 CTGCTAGGGATGCCATTGGCAGAGCTAATCCAGTGTGAAAGCGGACAA 673
Qy      696 TTT 698
Db      674 TCT 676
```

RESULT 11

```
US-10-425-114-25774
; Sequence 25774, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tadaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improv
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 25774
; LENGTH: 932
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3959-015-A1_FLI
US-10-425-114-25774
```

Query Match 12.8%; Score 102.6; DB 7; Length 9.

Best Local Similarity 52.9%; Pred. No. 7.4e-24;

Matches 337; Conservative 0; Mismatches 264; Indels 4;

```
Qy      82 TTTTCATACGCTTCTGATTAAGAAAGATACACCTTTATGCTGTAA 141
Db      136 TGTGACTACTATTTCTGTATATTGATCACTGTGTTTCGATATATTCAGTT 195
Qy      142 CGAGGTGTGTCTCAGAGACGACTCGCAGGCCCTGAGGGTCCAGTT 201
Db      196 GGAGC-----ACAAGAGGCTGGGTGCCAGGCCCCGAGGAGCAATC 249
Qy      202 CCTTTGCGGCTTTCTTGGCAGCCCAAGCTACATGAGGTGTGCTGAG 261
Db      250 TAACTGTGCTTTCTTGGCAGCGCGGCGACCATGAACATGTGCTCAAG 309
Qy      262 GACAGTCATGCAAGGAGATGACGGCTTTAGCTGAGCAAGCACTGAC 321
Db      310 GATGATTAACGAAGCAGATACAGGCCAAGCTGGCTGC----- 355
Qy      322 GACATCTGCAACAGCTGTGCTGTTTCAGCCCCCGCTCCTGTACATGAG 381
```

```

Db      356 GACAGCATGCTGAACGGCAGCGACGCCGT-----CATGAGCCGGTTGTTG 401
Qy      382 ATGCGAGGTTGAGAGAACATATGTTGCGGCATCAATCTTCCAGTTATCAACAGACT 441
Db      402 CTGGCAGCAACACGGTAGTAGCTGTGTCCAGATTGATGCAACAAAGAAC--GTGCA 459
Qy      442 GGTAAACCCCGCTGACAGTCCCGCTCAGGCAAGTGTCTATGCGAGCTCCCTCTAG 501
Db      460 GCAGCCCGCTGATGTTGCGGACCCAGCCAGGGGGTGGCGGCGATCTCCAAAGGGGGAA 519
Qy      502 ACCCGAGCCCAATCGATGCGGATCTTTCAGAGAACGCTGTGATGATGACAGATTAAAGT 561
Db      520 GGTAGGGCCGAACCGGTGACGGCCCTGCAAGAAAGGGTTGACTTACGGGATTCAACG 579
Qy      582 TGGCTGTGCAACCTTACTGCGCTTTTACATGGGTACTGGAACAACACACTTGCACATA 621
Db      580 CCGGTGTGGAACTTGTACTGCGCACTCCGACTCCGACCAAGACAGACTGCAAGTT 639
Qy      622 TGACTACAAAGCCGACGAGGCGAAGGCGATTGCGAAAGCTATCTCTGTGTCGCGCA 681
Db      640 CCACTACCGGACTGCTGCGCAGGAGCCCAATGCGAAAGCTAAATCCGGTGTGAAGCAGA 699
Qy      682 GAAGGTTGTCAAGTTTGTATGAGCATCCGTTAAGCTT 718
Db      700 CAAGCTCGACAAAGATCTAGGGGGGTTCCCTACGGTT 736

```

```

RESULT 12
US-10-425-115-71464
; Sequence 71464, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 71464
; LENGTH: 1114
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_16517C.1
US-10-425-115-71464

```

```

Query Match      12.8%; Score 102.6; DB 8; Length 1114;
Best Local Similarity 52.9%; Pred. No. 8.1e-24;
Matches 337; Conservative 0; Mismatches 264; Indels 36; Gaps 4;

Qy      82 TTTCTCATACGCTTCTGATTAGAAAGATACACCTTATTAATCTGTTAAAGATGCGC 141
Db      136 TGTCACTACTATATTTGATGATGATGATGTTTGTATTAATTAAGTAAGGACCT 195
Qy      142 CGAGCGTGTGTCAGAGACGACTTCGACAGCCCTGAGGGTCCAGTTATGTGCAAGAA 201
Db      196 GGAGC-----ACAAGAGGCTGGGTGCAAGGCCCCCAAGGAAACCACTCTCTGATCAA 249
Qy      202 CTTTTCGCGCTTCTTTCGAGGCAAGACTACATGAGGGTTGCTGGAAGTGTCAAGG 261
Db      250 TAACTGTGCTTCTTCGCGACGCGCGACATGAAATGTCTCCAAGTGCACAAAGAA 309
Qy      262 GACAGTCATGCAAGGAGATGATGAGGCTTTAGTCAGCAAGCACTACAGGCTCTCAGG 321
Db      310 GATGATTAACAAAGCAGATCAAGCCCAAGCTGCTC-----CTCTCTATTC 355
Qy      322 GACATCTGCAACAGCTGTGCTGTTCAAGCCCGCTCTCTGATGATGAGAACCAAGCTCAC 381
Db      356 GACAGCATGCTGAACGGCAGCGACGCCGT-----CATGAGCCGGTTGTTG 401

```

```

Qy      382 ATGCGAGGTTGAGAGAACATATGTTGCGGCATCAATCTTCCAGTTAT 441
Db      402 CTGGCAGCAACACGGTAGTAGCTGTGTCCAGATTGATGCAACAAAGAA 459
Qy      442 GGTAAACCCCGCTGACAGTCCCGCTCAGGCAAGTGTCTATGCGAG 501
Db      460 GCAGCCCGCTGATGTTGCGGACCCAGCCAGGGGGTGGCGGCGATCTC 519
Qy      502 ACCCGAGCCCAATCGATGCGGATCTTTCAGAGAACGCTGTGATGATGAC 561
Db      520 GGTAGGGCCGAACCGGTGACGGCCCTGCAAGAAAGGGTTGACTTACG 579
Qy      582 TGGCTGTGCAACCTTACTGCGCTTTTACATGGGTACTGGAACAACAC 621
Db      580 CCGGTGTGGAACTTGTACTGCGCACTCCGACTCCGACTTCCGACCAAGC 639
Qy      622 TGACTACAAAGCCGACGAGGCGAAGGCGATTGCGAAAGCTATCTCT 681
Db      640 CCACTACCGGACTGCTGCGCAGGAGCCCAATGCGAAAGCTAAATCCGG 699
Qy      682 GAAGGTTGTCAAGTTTGTATGAGCATCCGTTAAGCTT 718
Db      700 CAAGCTCGACAAAGATCTAGGGGGGTTCCCTACGGTT 736

```

```

RESULT 13
US-10-424-599-61443
; Sequence 61443, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Mol-
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Impro-
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 61443
; LENGTH: 1191
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_26497C.1
US-10-424-599-61443

```

```

Query Match      12.7%; Score 102.4; DB 7; Length 1
Best Local Similarity 70.8%; Pred. No. 9.8e-24;
Matches 136; Conservative 0; Mismatches 56; Indels 0;

Qy      509 CCCATCGATGCGGATCTTTCAGAGAGAGGCTGTGATTTGACAGATTTT 568
Db      617 CCAAAACGCTCAGAGCTGCAACAGGAGTGTGTTGACAGGCTTTT 676
Qy      569 GGCACCTCTACTGCGCTTTTACATCGTACTCGGACAAACACACTTGC 628
Db      677 GGTGACCTTTTGTGCTGTAATCTGCTACTACAGCAAGCAATATGCG 736
Qy      679 AAAGCCGACGAGGAGAAAGCATTTGCGAAAGCTTAATCTTGTGTG 688
Db      737 CGCACTGCTGCTCGGGATGCAATAGCTTAAACCAACCAAGTTGCAAG 796
Qy      689 GTCAAGTTTGA 700
Db      797 GATTAAGATCTAA 808

```

```

RESULT 14
US-10-425-114-27574
; Sequence 27574, Application US/10425114

```



```
Publication No. US2004003488A1
GENERAL INFORMATION:
APPLICANT: Liu, Jingdong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E
APPLICANT: Tabaska, Jack E
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(5313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 27574
LENGTH: 1032
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: LIB4732-053-A11_FLI
US-10-425-114-27574

Query Match      12.7%; Score 102.2; DB 7; Length 1032;
Best Local Similarity 51.2%; Pred. No. 1.1e-23;
Matches 288; Conservative 0; Mismatches 248; Indels 27; Gaps 1;

OY 156 AGGAGACGACCTCGAGGCCCCCTGAGGCTTATGTGCAAGAACCTTGGCGCTTCT 215
    |||||
DB 135 AGGAGGCTGGGTGCGAGGCCCCCGAGGAGCCCATCTCTGCAATCAATATCTGTGCTTCT 194
    |||||
OY 216 TCGGACGCAAGCTACATGAGGGTGTGCTCGAATGTGCTACCGAGAGACATGATCAAG 275
    |||||
DB 195 TCGGACGCGCGCGACATGAAATGTGCTCTAATGTGCAAGAGATGATTAAGAAC 254
    |||||
OY 276 CGAAGATGACGGCTTTAGCTGAGCAAGCCACTCAGGCTGCTCAGCGCATGTGCAAG 335
    |||||
DB 285 AGGA-----TCAGGCCAAGCTGGCTGCTCTCTATTTG 287
    |||||
OY 336 CTGCTGCTGTTCAAGCCCCCGCTCTGTATCATAGAACCAAGCTCATGCGAGTTGAGA 395
    |||||
DB 288 ACAGCATGTTGAACGAGCGAGCGCGCTCATGAGACCGGTGTGCTGCGAGCAACAGCG 347
    |||||
OY 396 GAACATGATTTGTGCGCGCATCATCTTCCAGCTATCAACAAGACCTGGTTACCCCGCTG 455
    |||||
DB 348 TAGTGTGCTGTGCCAAGTCGATGCAAAACAATGAACGTGAGACGCCCGCTGATTTTG 407
    |||||
OY 456 CAGTGGCCCTCAGGCAAGTAAATCTTATGCAAGCTCTCTTACGCCAGCCCAATC 515
    |||||
DB 408 CCGGACCCAGCGAGGGGGTGGCGCGCGATCTCAAAAGGGGGAAGGTAGGGCCGAAC 467
    |||||
OY 516 GATGGGATCTTTGCAAGAACGCTGTGATGACAGAAATTTAAGTGTGCGTGTGCAACC 575
    |||||
DB 468 GGTGAGGCGCTGCAAGAAAGGGTGGCTCAAGGATTCAACTGCGGTGTGGAACCT 527
    |||||
OY 576 TCTACTGCGCTTTACATCGTGTACTCGACAAACACACTTGCATATGACTTCAAAAGCG 635
    |||||
DB 528 TGTATGTGCGCTTCAACGCTACTCGACAAACAGACTGCAAGTTGCACTACCGGACTG 587
    |||||
OY 636 CAGGCAAGAAAGCAATTTGCAAAAGCTAATCTCTTGTGCTGCGCAAGAAAGTTGCAAGT 695
    |||||
DB 588 CTGCCAGGAGCGCATTCGTCGAAGGCTAATCCGTTGTGAGAGCAAGCAAGCTCGACAAGA 647
    |||||
OY 696 TTTGATGAGCATCCCTTAAGCTT 718
    |||||
DB 648 TCTAGGGGGGTTCCCTACGGTT 670
    |||||
```

RESULT 15
US-10-425-114-22047

; Sequence 22047, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong

```
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E
APPLICANT: Tabaska, Jack E
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules
TITLE OF INVENTION: Plants and Uses Thereof for Plant Impro
FILE REFERENCE: 38-21(5313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 22047
LENGTH: 736
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: LIB3356-023-F11_FLI
US-10-425-114-22047
```

Query Match 12.7%; Score 101.8; DB 7; Length 72
Best Local Similarity 53.5%; Pred. No. 1.2e-23;
Matches 301; Conservative 0; Mismatches 232; Indels

```
OY 156 AGGAGACGACCTCGAGGCCCCCTGAGGCTCAGTTATGTGCAAGAACCTT 215
    |||||
DB 114 AGGAGGCTGGGTGCGAGGCCCCCGAGGAGCCCATCTCTGCAATCAATTA 173
    |||||
OY 216 TCGGACGCAAGCTACATGAGGGTGTGCTCGAATGTGCTACCGAGAGAC 275
    |||||
DB 174 TCGGACGCGCGCGACATGAAATGTGCTCTAATGTGCAAGAGAT 233
    |||||
OY 276 CGAAGATGACGGCTTTAGCTGAGCAAGCCACTCAGGCTGCTCAGGCGAC 335
    |||||
DB 234 AGGATCAGGCCAAGCTGCTGCT-----CTCTCTATGAC 279
    |||||
OY 336 CTGCTGCTGTTCAAGCCCCCGCTCTGTATCATGAGAACCAAGCTCATAG 395
    |||||
DB 280 CCGGACGAGCGCGCT-----CATGAGCGGTTGTGCTG 325
    |||||
OY 396 GAACATGATTTGTGCGCGCATCATCTTCCAGCTATCAACAAGACCTGCT 455
    |||||
DB 326 GTAGTGTGTTGCTCCAAAGTTGAAGTTGCAAAACAATGAAC--GTGCAAGA 383
    |||||
OY 456 CAGTGGCCCTCAGGCAAGTAAATCTTATGCAAGCTCTCTTACGAC 515
    |||||
DB 384 TTGCTCGAACCGAGGAGGGGTGCGCGATCTCAAAAGGGGGAAGGT 443
    |||||
OY 516 GATGCGAATCTTTCAGGAAGCGTGTGATGACAGATTTAAGTGTG 575
    |||||
DB 444 GGTGAGGCGCTGCAAGAAAGGGTGGCTTACGAGGATTCAACTGCGCG 503
    |||||
OY 576 TCTACTGCGCTTTACATCGTGTACTCGACAAACACACTTGCATATGAG 635
    |||||
DB 504 TGTATGTGCACTCACCGCTACTCCGACAAAGCAAGACTGCAAGTTGCA 563
    |||||
OY 636 CAGGCAAGAAAGCAATTTGCAAAAGCTAATCTCTTGTGCTGCGCAAGAC 695
    |||||
DB 564 CTGCCAGGAGCGCATTCGTCGAAGGCTAATCCGTTGTGAGAGCAAGCAAG 623
    |||||
OY 696 TTTGATGAGCATCCCTTAAGCTT 718
    |||||
DB 624 TCTAGGGGGGTTCCCTACGGTT 646
    |||||
```

Search completed: December 8, 2005, 10:29:11
Job time : 803 secs


```

; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMT GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: PANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 47138
; LENGTH: 1067
; TYPE: DNA
; ORGANISM: Bovine 19866881619274
US-10-750-185-47138

Query Match      4.8%; Score 38.2; DB 6; Length 1067;
Best Local Similarity 54.7%; Pred. No. 0.0067;
Matches 76; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

Qy 1
567 GTGCAACCTTACTGCGCTTACTGCTGACTCGGACAAACACTGACATATGACT 626
574 GTGGAAATGTTACTGTAGTGTACACACTTACTCAGATGTACACACTGCTTCAATT 515
627 ACAAGCGCGACGAGGAGGAGGATTCGGAAGCTAATCTGTGCGGCGCGAAGG 686
514 AAGACTGATGCTGCTGTGAAAATCAGAAAAGAAAATCAGTACTGTTGGTGAAGA 455
687 TTGTCAAGTTTGTGATGAGC 705
454 TCCAGAGATTGGAAGCTGC 436
Db

RESULT 3
US-11-153-880-3/c
; Sequence 3, Application US/1153880
; Publication No. US20050256050A1
; GENERAL INFORMATION:
; APPLICANT: HU, JING-SHAN
; APPLICANT: ROSEN, CRAIG A.
; TITLE OF INVENTION: VASCULAR ENDOTHELIAL GROWTH FACTOR 2
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX
; STREET: 1100 NEW YORK AVENUE
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/11/153,880
; FILING DATE: 16-Jun-2005
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/219,442
; FILING DATE: 23-DEC-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/999,811
; FILING DATE:

```

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/465,968
; FILING DATE: 06-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: MARKOWICZ, KAREN R.
; REGISTRATION NUMBER: 36,351
; REFERENCE/DOCKET NUMBER: 1488.1000004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)371-2600
; TELEFAX: (202)371-2540
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1526 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 71..142
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 143..1120
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 71..1120
US-11-153-880-3

Query Match      4.1%; Score 32.8; DB 7; Length 15;
Best Local Similarity 61.9%; Pred. No. 0.6;
Matches 52; Conservative 0; Mismatches 32; Indels 0;

Qy
23 AGTGTGTGTCTTACTAGTGTGTTGCAAGCTGAAGCTGAGGAGAGA 82
484 AGTGTATTGGCAAAAGCATGTTGCTGTTGGGCGCTTGAGAGAGA 425
83 TTCTCATACGCTCTGATTAGGAA 106
424 TTCAATAGCGTCTTGCTGAGGTA 401
Db

RESULT 4
US-11-153-880-1/c
; Sequence 1, Application US/1153880
; Publication No. US20050256050A1
; GENERAL INFORMATION:
; APPLICANT: HU, JING-SHAN
; APPLICANT: ROSEN, CRAIG A.
; TITLE OF INVENTION: VASCULAR ENDOTHELIAL GROWTH FACTOR 2
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX
; STREET: 1100 NEW YORK AVENUE
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/11/153,880
; FILING DATE: 16-Jun-2005
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/219,442
; FILING DATE: 23-DEC-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/999,811
; FILING DATE:

```

FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/465,968
FILING DATE: 06-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: MARKOWICZ, KAREN R.
REGISTRATION NUMBER: 36,351
REFERENCE/DOCKET NUMBER: 1488.1000004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)371-2600
TELEFAX: (202)371-2540
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1674 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 12..80
FEATURE:
NAME/KEY: mac_peptide
LOCATION: 81..1268
FEATURE:
NAME/KEY: CDS
LOCATION: 12..1268
US-11-153-880-1

Query Match 4.1%; Score 32.8; DB 7; Length 1674;
Best Local Similarity 61.9%; Pred. No. 0.63; Mismatches 32; Indels 0; Gaps 0;

Matches 52; Conservative 0; Mismatches 32; Indels 0; Gaps 0;
QY 23 AGTGTGTGTCATTAGTGTGTCGAAGCTTGAGCGAGATTGCAAGATT 82
DB 632 AGTGTGATGGCAAACTGATGTTACTGTGTTGGGGCTTGAGAGAGGCACTGTAT 573
QY 83 TTCTCATACGCTTCGTGATTAGGA 106
DB 572 TTCAATATACGCTTGTGCTGAGGTA 549

RESULT 5
US-11-064-769-1/c
Sequence 1, Application US/11064769
GENERAL INFORMATION:
APPLICANT: Alltalo, Karl
APPLICANT: Hilunen, Mikko O
APPLICANT: Jeltsch, Markku M
APPLICANT: Achen, Marc G
TITLE OF INVENTION: Use of VEGF-C or VEGF-D Gene or Protein to Prevent Restenosis
FILE REFERENCE: 28967/35601A
CURRENT APPLICATION NUMBER: US/11/064,769
PRIOR FILING DATE: 2005-02-24
PRIOR APPLICATION NUMBER: US/09/427,657
PRIOR FILING DATE: 1999-10-26
PRIOR APPLICATION NUMBER: US 60/105,587
PRIOR FILING DATE: 1998-10-26
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 1997
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (352)..(1608)
US-11-064-769-1

Query Match 4.1%; Score 32.8; DB 7; Length 1997;
Best Local Similarity 61.9%; Pred. No. 0.7;

Matches 52; Conservative 0; Mismatches 32; Indels 0;
QY 23 AGTGTGTGTCATTAGTGTGTCGAAGCTTGAGCGAGATTGCAAGCTTGAGCGAG 82
DB 972 AGTGTGATGGCAAACTGATGTTACTGTGTTGGGGCTTGAGAGAGAG 913
QY 83 TTCTCATACGCTTCGTGATTAGGA 106
DB 912 TTCAATATACGCTTGTGCTGAGGTA 889

RESULT 6
US-11-064-774A-21/c
Sequence 21, Application US/11064774A
GENERAL INFORMATION:
APPLICANT: Alltalo et al.
TITLE OF INVENTION: MATERIALS AND METHODS INVOLVING HYBRID V
FILE REFERENCE: 28967/35977B2
CURRENT APPLICATION NUMBER: US/11/064,774A
PRIOR FILING DATE: 2005-02-24
PRIOR APPLICATION NUMBER: 09/795,006
PRIOR FILING DATE: 2001-02-26
PRIOR APPLICATION NUMBER: US 60/205,331
PRIOR FILING DATE: 2000-05-18
PRIOR APPLICATION NUMBER: US 60/185,205
PRIOR FILING DATE: 2000-02-25
NUMBER OF SEQ ID NOS: 1212
SOFTWARE: PatentIn Ver. 3.2
SEQ ID NO 21
LENGTH: 1997
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (352)..(1608)
US-11-064-774A-21

Query Match 4.1%; Score 32.8; DB 7; Length 1997;
Best Local Similarity 61.9%; Pred. No. 0.7; Mismatches 32; Indels 0;

Matches 52; Conservative 0; Mismatches 32; Indels 0;
QY 23 AGTGTGTGTCATTAGTGTGTCGAAGCTTGAGCGAGATTGCAAGCTTGAGCGAG 82
DB 972 AGTGTGATGGCAAACTGATGTTACTGTGTTGGGGCTTGAGAGAGAG 913
QY 83 TTCTCATACGCTTCGTGATTAGGA 106
DB 912 TTCAATATACGCTTGTGCTGAGGTA 889

RESULT 7
US-11-064-769-17/c
Sequence 17, Application US/11064769
GENERAL INFORMATION:
APPLICANT: Alltalo, Karl
APPLICANT: Hilunen, Mikko O
APPLICANT: Jeltsch, Markku M
APPLICANT: Achen, Marc G
TITLE OF INVENTION: Use of VEGF-C or VEGF-D Gene or Protein
FILE REFERENCE: 28967/35601A
CURRENT APPLICATION NUMBER: US/11/064,769
PRIOR FILING DATE: 2005-02-24
PRIOR APPLICATION NUMBER: US/09/427,657
PRIOR FILING DATE: 1999-10-26
PRIOR APPLICATION NUMBER: US 60/105,587
PRIOR FILING DATE: 1998-10-26
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 17

LENGTH: 2679
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: chimeric
OTHER INFORMATION: sequence in which CMV promoter sequence is ligated
OTHER INFORMATION: to Homo sapien VEGF-C sequence
US-11-064-769-17

Query Match 4.1%; Score 32.8; DB 7; Length 2679;
Best Local Similarity 61.9%; Pred. No. 0.83;
Matches 52; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 23 AGTGTGTGTCTATTAGTCTTGTGCAAGTCTGAAGCTTTAGCGAGATTGGCAGATT 82
DB 1654 AGTGTGATTGGCAAAACGATTGTACTGTGGGCTTTAGAGAGAGCGACTGTAT 1595
QY 83 TTCTCATACGCTTCTGATTAGGAA 106
DB 1594 TTCAATTAACGCTTCTGAGGTA 1571

RESULT 8
US-11-121-086-75/c
Sequence 75, Application US/11121086
Publication No. US20050266459A1
GENERAL INFORMATION:
APPLICANT: POULSEN, TIM S.
APPLICANT: NIELSEN, KIRSTEN V.
TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
FILE REFERENCE: 09138, 6000-00000
CURRENT APPLICATION NUMBER: US/11/121,086
CURRENT FILING DATE: 2005-05-04
PRIOR APPLICATION NUMBER: 60/567,570
PRIOR FILING DATE: 2004-05-04
NUMBER OF SEQ ID NOS: 107
SOFTWARE: PatentIn version 3.3
SEQ ID NO 75
LENGTH: 161874
TYPE: DNA
ORGANISM: Homo sapiens
US-11-121-086-75

Query Match 4.0%; Score 32; DB 7; Length 161874;
Best Local Similarity 50.0%; Pred. No. 17;
Matches 80; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 525 CTTGAGGAGCGGTGTGATTGACAGATTAAAGTGTGCGCAACTCTACTGCG 584
DB 131884 CTGACATGAAGCGTTGGGATTGTGATTAGTAGTTTCAGGGTGAGCACACAGAACTTG 131825
QY 565 CTTTACATCGTACTCGGCAACAACACTTGCACATAGACTACAAAGCCGAGGGCAGG 644
DB 131824 CTTATGATCATGACGTAACAGACATACCACTAGAGAAATACGAAAGTTGAGGTAGAG 131765
QY 645 AAGCGATTCCGAAGCTATCTCTTGTGTGCGCGAGAA 684
DB 131764 AGACAAGTAGAAGCTTATGCTCATCATGTCAGGTAGAA 131725

RESULT 9
US-11-121-086-20/c
Sequence 20, Application US/11121086
Publication No. US20050266459A1
GENERAL INFORMATION:
APPLICANT: POULSEN, TIM S.
APPLICANT: NIELSEN, KIRSTEN V.
TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
FILE REFERENCE: 09138, 6000-00000
CURRENT APPLICATION NUMBER: US/11/121,086
CURRENT FILING DATE: 2005-05-04
PRIOR APPLICATION NUMBER: 60/567,570
PRIOR FILING DATE: 2004-05-04

NUMBER OF SEQ ID NOS: 107
SOFTWARE: PatentIn version 3.3
SEQ ID NO 20
LENGTH: 162289
TYPE: DNA
ORGANISM: Homo sapiens
US-11-121-086-20

Query Match 3.9%; Score 31.6; DB 7; Length 16.
Best Local Similarity 49.1%; Pred. No. 23;
Matches 112; Conservative 0; Mismatches 114; Indels 1;

QY 23 AGTGTGTGTCTATTAGTGTGTTGCAAGTCTGAAGCCTTGAAGGAGA 82
DB 91649 AGAGTACTGTCTTGTCTGACTGACCTTGTGTGAATGACTTAAAGAGC 91590
QY 83 TTCTCATACGCTTCTGATTAGGAAAGATACACCTTATTAATCTGTAA 142
DB 91589 CTCTGCCATGGCTCTGGCTAAGTGGAGGCCCGTGTGGTCTGTGAC 91530
QY 143 GAGCGTGTCT--TCAGAGACGACCTTGCAGGCCCTCGAAGGTTCACT 200
DB 91529 CAGAGCTGTCTTGTGACTTGAAGGCTGAGGCTCCGACAGCGCAGC 91470
QY 201 ACCTTGGCGCTTCTTGTGCGACCAAGCTACCATGAGGTTGTGCTGA ;
DB 91469 GCTTCAGTCTCTCATCTGTGACAAAATTAACAGGCGCTGTGCTCCA ;

RESULT 10
US-11-121-086-35
Sequence 35, Application US/11121086
Publication No. US20050266459A1
GENERAL INFORMATION:
APPLICANT: POULSEN, TIM S.
APPLICANT: NIELSEN, KIRSTEN V.
TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
FILE REFERENCE: 09138, 6000-00000
CURRENT APPLICATION NUMBER: US/11/121,086
CURRENT FILING DATE: 2005-05-04
PRIOR APPLICATION NUMBER: 60/567,570
PRIOR FILING DATE: 2004-05-04
NUMBER OF SEQ ID NOS: 107
SOFTWARE: PatentIn version 3.3
SEQ ID NO 35
LENGTH: 170995
TYPE: DNA
ORGANISM: Homo sapiens
US-11-121-086-35

Query Match 3.9%; Score 31.2; DB 7; Length 17.
Best Local Similarity 50.7%; Pred. No. 33;
Matches 75; Conservative 0; Mismatches 73; Indels 0;

QY 152 TCTGAGAGAGCACTTCGACAGCCCTGAGGGTCCAGTTATGTGACAGA 211
DB 32928 TCTCTGACACCCAGCTCCCAACCAAGCAAGGGGCGCTCACTCGACA 32987
QY 212 TTCTTGGCAGCAAGCTACATGAGGTTGTGCTGGAAGTCTACCGAG 271
DB 32988 TTCCGGGGGGCGCCGAAGGCCCAAGGGGTCTGGGGGCACTTCAATTGC 33047
QY 272 CAAGGGAAGATGACGGCTTTAGCTGAGC 239
DB 33048 GAGCGGAGCTGAGGTGCTCGCGGAGC 33075

RESULT 11
US-10-750-185-46945/c
Sequence 46945, Application US/10750185
Publication No. US20050260603A1
GENERAL INFORMATION:
APPLICANT: MMI GENOMICS, INC.

APPLICANT: DENISE, Sue K.
APPLICANT: KERR, Richard
APPLICANT: ROSENFELD, David
APPLICANT: HOLM, Tom
APPLICANT: BATES, Stephen
APPLICANT: FANTIN, Dennis
FILE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
FILE REFERENCE: MM1100-2
CURRENT APPLICATION NUMBER: US/10/750,185
CURRENT FILING DATE: 2003-12-31
PRIOR APPLICATION NUMBER: US 60/437,482
PRIOR FILING DATE: 2002-12-31
NUMBER OF SEQ ID NOS: 64922
SOFTWARE: PatentIn version 3.1
SEQ ID NO: 46945
LENGTH: 2070
TYPE: DNA
ORGANISM: Bovine 19866880571876
US-10-750-185-46945

Query Match 3.9%; Score 31; DB 6; Length 2070;
Best Local Similarity 62.0%; Pred. No. 3;
Matches 49; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY 350 CCCCCGCTCTGTACATGAGCCAGCTCAGATGCGAGTTGAGAGACATGATGTC 409
DB 2050 CCCCCTATCCCATCATAGATTAACCTTGACATCCACCTGACAGACAGTATTG 1991
410 CCGCATCATCTTCCAGCT 428
DB 1990 AGACATTAGTTGCCATCT 1972

RESULT 12
US-11-121-086-58
Sequence 58, Application US/11121086
Publication No. US20050266459A1
GENERAL INFORMATION:
APPLICANT: POULSEN, TIM S.
APPLICANT: NIELSEN, KIRSTEN V.
FILE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
FILE REFERENCE: 09138.6000-00000
CURRENT APPLICATION NUMBER: US/11/121,086
CURRENT FILING DATE: 2005-05-04
PRIOR APPLICATION NUMBER: 60/567,570
PRIOR FILING DATE: 2004-05-04
NUMBER OF SEQ ID NOS: 107
SOFTWARE: PatentIn version 3.3
SEQ ID NO: 58
LENGTH: 180654
TYPE: DNA
ORGANISM: Homo sapiens
US-11-121-086-58

Query Match 3.9%; Score 31; DB 7; Length 180654;
Best Local Similarity 49.7%; Pred. No. 40;
Matches 79; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 587 TTACATCGCTACCGACAAACACACTTGACATATGACTACAAAGCCGACGAGGAA 646
DB 29799 TTTCCACAAATATGGGGCTAACTCTCCCTCAATGACTTGAAGCCGAGGAAATTA 29858
QY 647 GCGATTGGAAGCTAATCTCTTCTGCTGCGCGAAGGTTGCAAGTTTGTATGAGCA 706
DB 29859 TGGATTACTATATTAACCTTCTTACCTGTCTATATAAGGCAATTAAGAACTGGA 29918
QY 707 TCCGTTAAGCTTTGCGCGAGAGATTAGGCTTCATACA 745
DB 29919 TCATTAAAGCTGTATCTTCATTAATTGGTCACATCA 29957

RESULT 13
US-10-750-185-54345/c

Sequence 54345, Application US/10750185
Publication No. US20050260603A1
GENERAL INFORMATION:
APPLICANT: MMI GENOMICS, INC.
APPLICANT: DENISE, Sue K.
APPLICANT: KERR, Richard
APPLICANT: ROSENFELD, David
APPLICANT: HOLM, Tom
APPLICANT: BATES, Stephen
APPLICANT: FANTIN, Dennis
FILE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
FILE REFERENCE: MM1100-2
CURRENT APPLICATION NUMBER: US/10/750,185
CURRENT FILING DATE: 2003-12-31
PRIOR APPLICATION NUMBER: US 60/437,482
PRIOR FILING DATE: 2002-12-31
NUMBER OF SEQ ID NOS: 64922
SOFTWARE: PatentIn version 3.1
SEQ ID NO: 54345
LENGTH: 1433
TYPE: DNA
ORGANISM: Bovine 19866880647140
US-10-750-185-54345

Query Match 3.8%; Score 30.2; DB 6; Length 14
Best Local Similarity 55.1%; Pred. No. 4.6;
Matches 59; Conservative 0; Mismatches 48; Indels 0;

QY 657 AAGTAATCTCTTGTGCGCCGAGAGGTTGCAAGTTTGTATGAGC 716
DB 828 AAGTAATCATGAAGTTGATGCTGATTAAGACAGTTGTTTGAAGACT 769
QY 717 TTTTCTGCCGAGATTAGGCTTCATACATGAGTACTTACATCT
DB 768 TTTCTGGAACATAATTGGTCATAAAGCTGAATATAGCATGCTCT

RESULT 14
US-10-750-185-33374
Sequence 33374, Application US/10750185
Publication No. US20050260603A1
GENERAL INFORMATION:
APPLICANT: MMI GENOMICS, INC.
APPLICANT: DENISE, Sue K.
APPLICANT: KERR, Richard
APPLICANT: ROSENFELD, David
APPLICANT: HOLM, Tom
APPLICANT: BATES, Stephen
APPLICANT: FANTIN, Dennis
FILE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
FILE REFERENCE: MM1100-2
CURRENT APPLICATION NUMBER: US/10/750,185
CURRENT FILING DATE: 2003-12-31
PRIOR APPLICATION NUMBER: US 60/437,482
PRIOR FILING DATE: 2002-12-31
NUMBER OF SEQ ID NOS: 64922
SOFTWARE: PatentIn version 3.1
SEQ ID NO: 33374
LENGTH: 1523
TYPE: DNA
ORGANISM: Bovine 19866881033634
US-10-750-185-33374

Query Match 3.7%; Score 30; DB 6; Length 1523;
Best Local Similarity 54.5%; Pred. No. 5.5;
Matches 60; Conservative 0; Mismatches 50; Indels 0;

QY 122 AGTCTGTTAAGATGCCACGAGCGCTGTCTCAGAGACGACCTGC 181
DB 889 AGTGTATGTGAGATGCTCTGAGAGTGAAGGACCAATGAGAGGCTGCTG 948
QY 182 GGTCCAGTTATGTGCAAGACCTTTGGCGCTTCTTCCGACGCAAGCTA

Db 949 ATTCTTCTTTGTGCATCATATTATGGCTTGCAGATGGTCAATGCTCC 998

RESULT 15

US-10-750-185-56847
; Sequence 56847, Application US/10750185
; Publication No. US20050260603A1
GENERAL INFORMATION:
; APPLICANT: MMT GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: PANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: Patent version 3.1
; SEQ ID NO 56847
; LENGTH: 837
; TYPE: DNA
; ORGANISM: Bovine 19866880892313
US-10-750-185-56847

Query Match 3.7%; Score 29.8; DB 6; Length 837;
Best Local Similarity 54.0%; Pred. No. 4.6;
Matches 61; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

OY 502 ACCGAGCCCAATCGATCGGATCTTGACAGAGCGTGTGATTGACAGATTAAAGTG 561
Db 450 ACGGGGGCCCATTTCTGTGACTGTGTGGCAGGAAATGCTTCTGTGAAAGATTGACAGGG 509
OY 562 TCGCTGTGGCAACCTCTACTGCGCTTACATCGTACTCGACAAACACACTT 614
Db 510 CCAGTGGGGAAGCTAAGCAGCCCTTGATTTCTTTCTTGCGTGGCAACTT 562

Search completed: December 8, 2005, 10:33:29
Job time : 250 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 2005, 12:54:20 ; Search time 134 Seconds
(without alignment)
616,442 Million cell updates/sec

Title: US-10-716-089-20

Perfect score: 984
Sequence: 1 MATEVSGEETTSQARPECPV.....GGEALAKANPLVAVKVKYKF 188

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- A_Geneseq.21:*
- 1: geneseqp1980s:*
 - 2: geneseqp1990s:*
 - 3: geneseqp2000s:*
 - 4: geneseqp2001s:*
 - 5: geneseqp2002s:*
 - 6: geneseqp2003as:*
 - 7: geneseqp2003bs:*
 - 8: geneseqp2004s:*
 - 9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description.
1	984	100.0	188	5	ABG93809 Transcrip
2	984	100.0	188	5	AAU71800 Zinc fing
3	416.5	42.3	173	7	ABM90068 Rice abio
4	416.5	42.3	173	7	ABM87792 Rice abio
5	412.5	41.9	179	8	ADY97065 Plant ful
6	412.5	41.9	179	8	ADY09608 Plant ful
7	408.5	41.5	169	7	ABM73978 DNA clone
8	402.5	40.9	173	8	ADY75512 Plant ful
9	402.5	40.9	173	8	ADY94337 Plant ful
10	401.5	40.8	173	8	ADY5012 Plant ful
11	386	39.2	160	7	ABM90052 Rice abio
12	386	39.2	169	7	ABM87359 Rice abio
13	379	38.5	173	3	AAQ32536 Arabidops
14	379	38.5	173	3	AAQ49664 Arabidops
15	379	38.5	173	3	AAQ49652 Arabidops
16	378	38.4	173	3	AAQ23282 Arabidops
17	374	38.0	173	3	AAQ54935 Arabidops
18	366	37.2	213	2	AA422772 Rat neuro
19	365	37.1	213	2	ABU52627 Human bra
20	365	37.1	229	5	ABP41626 Human ova
21	365	37.1	310	3	AAQ56766 Human pro
22	364	37.0	213	4	AAQ64505 Gene 24 h
23	364	37.0	213	4	AAQ98228 Human zin
24	364	37.0	213	7	ABG75067 Human zin

25	364	37.0	213	8	AD124504	Ad1:
26	364	37.0	213	8	ADN04776	Adn:
27	364	37.0	213	8	ADP24800	Adp:
28	364	37.0	213	9	ADP07989	Adx:
29	364	37.0	213	9	AEA98119	Aea:
30	362.5	36.8	170	3	AAQ14725	Aag:
31	361.5	36.7	175	3	AAQ43065	Aag:
32	361.5	36.7	175	3	AAQ17575	Aag:
33	359.5	36.5	170	3	AAQ17628	Aag:
34	359	36.5	190	7	ABM86938	Abm:
35	357	36.3	224	8	ADY23911	Ady:
36	356	36.2	161	3	AAQ43390	Aag:
37	356	36.2	161	3	AAQ09127	Aag:
38	355	36.1	233	8	ADP22660	Adp:
39	352	35.8	223	7	ADP59857	Adp:
40	351.5	35.7	170	3	AAQ16569	Aag:
41	351	35.7	205	8	ADP22662	Adp:
42	346.5	35.2	163	3	AAQ30169	Aag:
43	346.5	35.2	163	3	AAQ49841	Aag:
44	346.5	35.2	163	3	AAQ49016	Aag:
45	346.5	35.2	199	3	AAQ49940	Aag:

ALIGNMENTS

RESULT 1	ABG93809	standard; protein; 188 AA.
ID	ABG93809	
XX	ABG93809;	
DT	28-NOV-2002	(first entry)
XX		
DE	Transcription factor stress-related protein (TFSRP) #4.	
XX		
KW	Transcription factor stress-related protein; TFSRP; MYB-1	
KW	AP2 similar-2 protein; APS-2; zinc-finger factor-2; ZF-2;	
KW	ZF-4 protein; ZF-5 protein; CAA1-Box binding factor prote:	
KW	Sigma factor like protein; SFL-1; transgenic; plant; drouc	
KW	environmental stress; Physcomitrella patens; high salinity	
KW	ciliate; fungus.	
XX		
OS	Physcomitrella patens.	
XX		
PN	US2002102695-A1.	
XX		
PD	01-AUG-2002.	
XX		
PF	06-APR-2001; 2001US-00828303.	
XX		
PR	07-APR-2000; 2000US-0196001P.	
XX		
PA	(SILV/) SILVA O D C E.	
PA	(BOHN/) BOHNERT H J.	
PA	(THIE/) THIELEN N V.	
PA	(CHEN/) CHEN R.	
XX		
PI	Silva ODE, Bohnert HJ, Thielen NV, Chen R;	
XX		
DR	WPI; 2002-690614/74.	
XX		
DR	N-PSDB; ABS70622.	
XX		
PT	Novel transcription factor stress-related protein such as	
PT	protein, zinc-finger protein, or CAA1-Box binding factor f	
PT	useful for modifying stress tolerance of plant.	
XX		
PS	Claim 2; Fig 3D; 104dp; English.	
XX		
CC	The invention relates to an isolated transcription factor	
CC	protein (TFSRP) (I) selected from AP2 similar-2 protein (A	
CC	finger factor-2 (ZF-2) protein, ZF-3 protein, ZF-4 protein	
CC	MYB-1 protein, CAA1-Box binding factor protein-3 (CABP-3),	
CC		

CC like protein (SFL-1) and their orthologues. Also described is: (1) an
CC isolated TFSRP coding nucleic acid (II) coding for (I); (2) a transgenic
CC plant cell (III) transformed by (II), where the expression of (II) in the
CC plant cell results in increased tolerance to an environmental stress as
CC compared to a wild-type variety of the plant cell; (3) a transgenic plant
CC (IV) comprising (III); (4) a seed (V) produced by (IV), where the seed is
CC true breeding for an increased tolerance to environmental stress as
CC compared to a wild-type variety of the plant cell; (5) an agricultural
CC product (VI) produced by (IV) or (V); and (6) an isolated recombinant
CC expression vector (VII) comprising (II), where expression of (VII) in a
CC host cell results in increased tolerance to environmental stress as
CC compared to a wild-type variety of the host cell. (II) is useful for
CC modifying (increasing or decreasing) stress tolerance of a plant
CC (transgenic or not transgenic), by modifying the expression of TFSRP in
CC the plant. The plant is transformed with a promoter that directs
CC expression of the TFSRP. The promoter is tissue specific and is
CC developmentally regulated. TFSRP expression is modified by administration
CC of an antisense molecule that inhibits expression of TFSRP. (VII) is
CC useful for producing a transgenic plant containing (II), where expression
CC of the nucleic acid in the plant results in increased tolerance to
CC environmental stress as compared to a wild-type variety of the plant, by
CC generating a plant cell with (VII) comprising the nucleic acid,
CC tolerance to environmental stress as compared to a wild-type variety of
CC the plant. (I) is useful for conferring stress tolerance such as drought,
CC cold and/or salt tolerance to plants. (I), (II), (III) or (VII) is useful
CC for identifying Physcomitrella patens and related organisms, for mapping
CC of genomes of organisms related to P.patens, for identifying and
CC localizing P.patens sequences of interest, for evolutionary studies, for
CC determining TFSRP regions required for function, for modulating TFSRP
CC activity, for modulating metabolism of one or more cell functions, for
CC modulating transmembrane transport of one or more compounds, and for
CC modulating stress resistance. (I) is useful for reducing stress tolerance
CC response plants or more particularly, in the transcription of a protein
CC involved in a stress tolerance response in a P.patens plant. (II) is
CC useful for transforming plants and thus inducing tolerance to stresses
CC such as drought, high salinity and cold, for identifying the presence of
CC P.patens or a related organism in a mixed population of microorganism,
CC serve as marker for specific regions of the genome for mapping the genome
CC and for the functional studies of P.patens proteins. (II) is useful as
CC reference points for mapping the moss genomes, or of genomes of related
CC organisms, for evolutionary and protein structural studies, for
CC generating knockout mutation in the genomes of various organisms such as
CC bacteria, mammalian cells, yeast cells and plant cells, useful for
CC evaluating their ability or capacity to tolerate various stresses
CC conditions and the effect on the phenotype and/or genotype of the
CC mutation. (II) is useful as marker for the construction of a genomic map
CC in related mosses. (I) or (II) is useful for generating algae, ciliates,
CC plants, fungi or other microorganisms expressing mutated TFSRP nucleic
CC acid and protein molecules for improving stress tolerance. ABG93806-
CC ABG93813 represent P. patens TFSRP amino acid sequences of the invention
XX
XX
SQ Sequence 188 AA;

Query Match 100.0%; Score 984; DB 5; Length 188; 1
Best Local Similarity 100.0%; Pred. No. 6.9e-92; Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WATERVSOETTSSOAPBEGPVMCKNLGFFGSOATMGLCSKCYRETYWQAKMTALAEQATCA 60
DB 1 MATEHVSOETTSSOAPBEGPVMCKNLGFFGSOATMGLCSKCYRETYWQAKMTALAEQATCA 60
QY 61 AQAATATAAANVPAPVHETKLTCEVERTMIVPHOSSSYQODLVTPAAAAPQAVKSSIAA 120
DB 61 AQAATATAAANVPAPVHETKLTCEVERTMIVPHOSSSYQODLVTPAAAAPQAVKSSIAA 120
QY 121 PSRPPNCGSGCRKRVGLTGFRCGNNLYCALHRYSDKHTCTYDVKKAQGEAIAANPLV 180
DB 121 PSRPPNCGSGCRKRVGLTGFRCGNNLYCALHRYSDKHTCTYDVKKAQGEAIAANPLV 180
QY 181 VAEKVVKF 188
DB 181 VAEKVVKF 188

RESULT 2
AAU71800
ID AAU71800 standard; protein, 188 AA.
AC AAU71800;
XX 26-FEB-2002 (first entry)
DE Zinc finger-4 (ZF-4).
KW Transcription factor stress-related protein; AP2 similar-2
KW AP2-2; zinc-finger factor protein; ZF-2; ZF-3; ZF-4; ZF-5;
KW MYB-1 MYB-1 protein, CAAT-box binding factor protein-3; CA
KW Sigma factor like protein; SFL-1; environmental stress; mo
XX Physcomitrella patens.
OS
XX WO200177311-A2.
XX
XX 18-OCT-2001.
XX
XX 06-APR-2001; 2001WO-US011393.
XX
XX 07-APR-2000; 2000US-0196001P.
XX
XX (BADI) BASF PLANT SCI GMBH.
XX
XX Da Costa Silva O, Bohnert HJ, Van Thiele N, Chen R;
XX MPI, 2002-049146/06.
XX N-PSDB; AAS95832.
XX
XX New polypeptide, useful for identification of Physcomitrella
XX for modulating stress resistance of a plant, comprises an
XX transcription factor stress-related protein derived from P
XX patens.
XX
XX Claim 13, Fig 3D, 129pp; English.
XX
XX
XX The invention relates to an isolated transcription factor;
XX protein (TFSRP) (I) selected from an AP2 similar-2 protein
XX zinc-finger factor-2 protein (ZF-2), ZF-3 protein, ZF-4 pr
XX protein, a MYB-1 protein (MYB-1), a CAAT-box binding facto
XX (CABF-3), a sigma factor like protein (SFL-1) and its orth
XX described is an isolated TFSRP coding nucleic acid (II) wh
XX (I); and an isolated recombinant expression vector (III) c
XX where expression of (III) in a host cell results in increa
XX to environmental stress as compared to a wild type variety
XX cell. (I) or (II) is useful as markers for specific region
XX genome. (I) or (II) is useful for identification of Physco
XX and related organisms, mapping of genomes of P. patens
XX patens, identification and localisation of P. patens sequ
XX interest, evolutionary studies, determination of TFSRP reg
XX for function, modulation of a TFSRP activity, modulation o
XX metabolism of one or more cell functions, modulation of th
XX transport of one or more compounds and modulation of stres
XX (II) is useful for generating probes and primers for ident
XX cloning TFSRP homologues in other cell types and other org
XX as TFSRP homologues from mosses and related species, and f
XX and protein structural studies. AAU71797-AAU71804 represen
XX TFSRP amino acid sequences of the invention
XX
XX
SQ Sequence 188 AA;

Query Match 100.0%; Score 984; DB 5; Length 18
Best Local Similarity 100.0%; Pred. No. 6.9e-92; Matches 188; Conservative 0; Mismatches 0; Indels 0;

QY 1 WATERVSOETTSSOAPBEGPVMCKNLGFFGSOATMGLCSKCYRETYWQAK 60
DB 1 MATEHVSOETTSSOAPBEGPVMCKNLGFFGSOATMGLCSKCYRETYWQAK 60

Best Local Similarity 42.7%; Pred. No. 4,8e-34;
Matches 82; Conservative 28; Mismatches 47; Indels 35; Gaps 3;
Qy 8 QETTSQAPBEGPVMCKNLGFGFSQATMGLCSKCYRETYMQAKMTALAQTA----- 60
Db 4 KEACGQOEGPILCINNCGFFGSAAATMNMCKCHEKEMIKERQAKLAASSIYNCGDG 63
Qy 61 -----AQATSATAAAYQPPAPVHETKLTCEVERTMIIPHQSSVQODLVTPAAAPQAVK 115
Db 64 GKEHTVAASGSTAVAN-----AQVEAKTLV-----VQPTDVAGTSEE 100
Qy 116 SSIAAPSRPEPNRCGSKRKGVLGTGFCRCGNLYCALHRYSDKHTCTYDYKAAQGEALAK 175
Db 101 VAVVKVKEGPNRCATCRKRVGLTGFNCRCGNMYCALHRYSDKHFCQPDYRTAAADALAK 160
Qy 176 ANPLVAEKVVK 187
Db 161 ANPVVKAEXLDD 172
RESULT 5
ADY97065
ID ADX97065 standard; protein, 179 AA.
AC ADX97065;
XX
DT 21-APR-2005 (first entry)
XX
DE Plant full length insert polypeptide seqid 59729.
XX
KW plant protectant; plant growth regulant; gene therapy; plant;
KW recombinant DNA construct; physical array; plant breeding marker;
KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KW extreme osmotic condition; pathogen tolerance; pest tolerance;
KW growth rate; cell cycle pathway; disease resistance;
KW galactomannan production; lignin production; plant growth regulator;
KW yield; plant growth; plant development; seed oil; protein yield;
KW protein content.
XX
OS Unidentified.
XX
PN US2004034888-A1.
XX
PD 19-FEB-2004.
XX
PF 28-APR-2003; 2003US-00425114.
XX
PR 06-MAY-1999; 99US-00304517.
PR 05-NOV-2001; 2001US-00985678.
XX
PA (LITU/) LITU J.
PA (ZHOU/) ZHOU Y.
PA (KOVA/) KOVALIC D K.
PA (SCRE/) SCREEN S E.
PA (TABAA/) TABASKA J E.
PA (CAOY/) CAO Y.
XX
PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
XX
DR WPI, 2004-180133/17.
XX
PT New recombinant DNA construct, useful for improving plant tolerance to
PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
PT pests, for conferring increased resistance to plant disease, or for
PT improving yield.
PS Claim 1; SEQ ID NO 59729; 15pp; English.
XX
CC The invention describes a recombinant DNA construct comprising a
CC polynucleotide consisting of a sequence encoding an amino acid sequence
CC available in electronic form from the US patent office at
CC ftp.segdata.nsls.gov/sequence.html?DocID:2004034888. The polynucleotide
CC of the invention are also useful in physical arrays of molecules and as

CC plant breeding markers. The recombinant DNA construct is u
CC improving plant tolerance to cold, heat, drought, herbicid
CC osmotic conditions, pathogens or pests, for manipulating g
CC plant cells by modification of the cell cycle pathway, for
CC increased resistance to plant disease, for producing galac
CC lignin or plant growth regulators, for increasing the rate
CC recombinant in plants, for improving yield by modificati
CC photosynthesis or carbohydrate, nitrogen or phosphorus use
CC or by providing improved plant growth and development unde
CC stress condition or for modifying seed oil or protein yield
CC content. This is the amino acid sequence of a plant full l
CC polypeptide that can be used in the recombinant DNA constr
CC invention.
XX
SQ Sequence 179 AA;
Qy Query Match 41.9%; Score 412.5; DB 8; Length 1
Best Local Similarity 42.3%; Pred. No. 1.3e-33;
Matches 80; Conservative 28; Mismatches 50; Indels
Qy 8 QETTSQAPBEGPVMCKNLGFGFSQATMGLCSKCYRETYMQAKMTALAAC 60
Db 12 KEACGQOEGPILCINNCGFFGSAAATMNMCKCHEKEMIKERQAKLAAS 71
Qy 61 --AQATSATAAAYQPPAPVHETKLTCEVERTMIIPHQSSVQODLVTPA 118
Db 72 GKGRPAIATVGVAVP-----QVEKTLA-----VQPK 109
Qy 119 AAPSRPEPNRCGSKRKGVLGTGFCRCGNLYCALHRYSDKHTCTYDYKA 178
Db 110 IAKKEGPNRCATCRKRVGLTGFNCRCGNMYCALHRYSDKHFCQPDYRT 169
Qy 179 LVVAEKVVK 187
Db 170 VVKAEXLDD 178
RESULT 6
ADY09608
ID ADY09608 standard; protein, 179 AA.
XX
AC ADY09608;
XX
DT 21-APR-2005 (first entry)
XX
DE Plant full length insert polypeptide seqid 65423.
XX
KW plant protectant; plant growth regulant; gene therapy; pla
KW recombinant DNA construct; physical array; plant breeding
KW cold tolerance; heat tolerance; drought tolerance; herbici
KW extreme osmotic condition; pathogen tolerance; pest tolera
KW growth rate; cell cycle pathway; disease resistance;
KW galactomannan production; lignin production; plant growth
KW yield; plant growth; plant development; seed oil; protein
KW protein content.
XX
OS Unidentified.
XX
PN US2004034888-A1.
XX
PD 19-FEB-2004.
XX
PF 28-APR-2003; 2003US-00425114.
XX
PR 06-MAY-1999; 99US-00304517.
PR 05-NOV-2001; 2001US-00985678.
XX
PA (LITU/) LITU J.
PA (ZHOU/) ZHOU Y.
PA (KOVA/) KOVALIC D K.
PA (SCRE/) SCREEN S E.
PA (TABAA/) TABASKA J E.
PA (CAOY/) CAO Y.

XX
PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
XX
DR WPI, 2004-180133/17.
XX
PT New recombinant DNA construct, useful for improving plant tolerance to
XX cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
PT pests, for conferring increased resistance to plant disease, or for
XX improving yield.
XX
PS Claim 1; SEQ ID NO 65423; 15bp; English.
XX
XX The invention describes a recombinant DNA construct comprising a
XX polynucleotide consisting of a sequence encoding an amino acid sequence
XX available in electronic form from the US patent office at
XX ftp.segdata.uspbo.gov/sequence.html?DocID:2004034889. The polynucleotide
XX of the invention are also useful in physical arrays of molecules and as
XX plant breeding markers. The recombinant DNA construct is useful for
XX improving plant tolerance to cold, heat, drought, herbicides, extreme
XX osmotic conditions, pathogens or pests, for manipulating growth rate in
XX plant cells by modification of the cell cycle pathway, for conferring
XX increased resistance to plant disease, for producing galactomannan,
XX lignin or plant growth regulators, for increasing the rate of homologous
XX recombination in plants, for improving yield by modification of
XX photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
XX or by providing improved plant growth and development under at least one
XX stress condition or for modifying seed oil or protein yield and/or
XX content. This is the amino acid sequence of a plant full length insert
XX polypeptide that can be used in the recombinant DNA construct of the
XX invention.
XX
SQ Sequence 179 AA;
XX
Query Match 41.9%; Score 412.5; DB 8; Length 179;
Best Local Similarity 42.3%; Pred. No. 1.3e-33;
Matches 80; Conservative 28; Mismatches 50; Indels 31; Gaps 3;
XX
QY 8 QETTGAPEPGPVMKCNLCGFFGSOATMGLCSKYRETVMAQAKTALAEATQA----- 60
DB 12 KEAGCGQRPGLCTINCCGFFGSAATMNCCKHCKEMIKOBQAOCLASISIVGCGDN 71
QY 61 --AQTATATTAATAVPPAPVHETKLTCEVERTMI VPHQSSYQODLTTPAAAAPQAVKSSI 118
DB 72 GNGPAAATVGA V-----QVEKRTIA-----VQPMVAETSSAAAV 109
QY 119 AAPSPPEPNRCGSKRKGVLTFPKRCGNYCALHRYSDKHTCTYDKAAGQEAIAKANP 178
DB 110 IAKAKGFPKRCATCRKRGVLTGPNCGCGNTYCSMHRYSKHKDQCPDYRRAADALAKANP 169
QY 179 LVVAEKVVK 187
DB 170 VVKAERLKD 178
XX
RESULT 7
ABM73978
ID ABM73978 standard; protein; 169 AA.
XX
AC ABM73978;
XX
DT 17-OCT-2003 (first entry)
XX
DE DNA clone ordinating in barley containing SNP sequence #388.
XX
XX Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis.
XX
OS Hordeum vulgare.
XX
PN WO2003057877-A1.
XX
PD 17-JUL-2003.
XX
PF 16-DEC-2002; 2002WO-1B005403.

XX
PR 20-DEC-2001; 2001JP-00387059.
XX
PR 20-DEC-2001; 2001JP-00387131.
XX
PR 20-DEC-2001; 2001JP-00403299.
XX
PR 20-DEC-2001; 2001JP-00403300.
XX
PR 27-SEP-2002; 2002JP-00327515.
XX
XX (UNII-) UNIV JAPAN OKAYAMA.
XX
XX Sato K, Takeda K, Kohara Y;
XX
DR WPI; 2003-587127/55.
XX
XX Single nucleotide polymorphism sites in barley varieties;
XX PT sequences containing them for analysis and identification
XX PT varieties and production of barley transformants with des-
XX characteristics.
XX
PS Disclosure; SEQ ID NO: 284dp; Japanese.
XX
XX The present invention relates to oligonucleotide clones of
XX barley (Hordeum vulgare) which contain single nucleotide
XX (SNP). The oligonucleotides may be used for analysis of SN
XX varieties, identification of particular varieties and gen
XX analysis, isolation of specific genes and creation of new
XX transformation of barley varieties with them and producti
XX varieties with desired properties. The present sequence is
XX oligonucleotide clone sequence featured in the specificat
XX sequence data for this patent did not form part of the pri
XX specification, but was obtained in electronic format direc
XX at ftp.wipo.int/pub/published-pct-sequences
XX
SQ Sequence 169 AA;
XX
Query Match 41.5%; Score 408.5; DB 7; Length 1
Best Local Similarity 45.3%; Pred. No. 3e-33;
Matches 86; Conservative 25; Mismatches 44; Indels 5;
XX
QY 8 QETTGAPEPGPVMKCNLCGFFGSOATMGLCSKYRETVMAQAKTALAEATQA----- 60
DB 4 KETGCGSREGPILCTVNCGFFGSAATMNCCKHCKEMIKOBQAOCLAAI 71
QY 58 TQAAATATTAATAVPPAPVHETKLTCEVERTMI VPHQSSYQODLVTI 117
DB 64 VKEHLAAGSTAVAV-----AHVQAKTTLTAQPA----DIAGI 106
QY 118 IAAPEPNRCGSKRKGVLTFPKRCGNYCALHRYSDKHTCTYDYI 177
DB 107 -----PSRSTKRAKRGVLTGPNCGCGNTYCATHRYSKHKCFDYI 158
QY 178 PLVVAEKVVK 187
DB 159 PVVKAERLKD 168
XX
RESULT 8
ADX75512
ID ADX75512 standard; protein; 173 AA.
XX
AC ADX75512;
XX
DT 21-APR-2005 (first entry)
XX
DE Plant full length insert polypeptide seqid 44878.
XX
XX plant protectant; plant growth regulant; gene therapy; pla
XX recombinant DNA construct; physical array; plant breeding
XX cold tolerance; heat tolerance; drought tolerance; herbic
XX extreme osmotic condition; pathogen tolerance; pest tolera
XX growth rate; cell cycle pathway; disease resistance;
XX galactomannan production; lignin production; plant growth
XX yield; plant growth; plant development; seed oil; protein
XX protein content.

Db 6 KEAGCQOPEGPILCINNCGFSGSAATMNMCSKCHKE-----MITKQEDAOQLAASSISDI 59
Qy 67 -----TAAAVQPPAPVHETKLTCEVERTMIVPHQSSSYQODLVTTPAAAPQAVK 115
Db 60 VNGDGGKGPVLAASVNAVVP-----QVEQKTIY-----VQPMVLVETSEA 100
Qy 116 SSIAAPSRPEPRRCGSKRVRGLTGPKRCGNYCALHRYSDKHTCTYTKAAGEALTK 175
Db 101 AAVTPKAKEGPDRCACRKRVLGTGFSRCGNNYCSVHRYSBKHCQDFYRTAARDALAK 160
Qy 176 ANPLVVAEKVVK 187
Db 161 ANPVVRAEKLDK 172
RESULT 10
ADX95012
ID ADX95012 standard; protein; 173 AA.
AC ADX95012;
DT 21-APR-2005 (first entry)
XX Plant full length insert polypeptide seqid 57676.
XX plant procectant; plant growth regulant; gene therapy; plant;
XX recombinant DNA construct; physical array; plant breeding marker;
XX cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
XX extreme osmotic condition; pathogen tolerance; pest tolerance;
XX growth rate; cell cycle pathway; disease resistance;
XX galactomannan production; lignin production; plant growth regulator;
XX yield; plant growth; plant development; seed oil; protein yield;
XX protein content.
XX Unidentified.
XX OS US2004034888-A1.
XX PN 19-FEB-2004.
XX PD 28-APR-2003; 2003US-00425114.
XX PF 06-MAY-1999; 99US-00304517.
XX PR 05-NOV-2001; 2001US-00985678.
XX (LIJU/) LIU J.
XX (ZHOU/) ZHOU Y.
XX (KOVAV/) KOVALIC D K.
XX (SCRE/) SCREEN S E.
XX (TABAK/) TABASKA J E.
XX (CAOY/) CAO Y.
XX
XX Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
XX WPI; 2004-180133/17.
XX
XX New recombinant DNA construct, useful for improving plant tolerance to
XX cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
XX pests, for conferring increased resistance to plant disease, or for
XX improving yield.
XX
XX Claim 1, SEQ ID NO 57676; 15pp; English.
XX
XX The invention describes a recombinant DNA construct comprising a
XX polynucleotide consisting of a sequence encoding an amino acid sequence
XX available in electronic form from the US patent office at
XX ftp.secdat.uspco.gov/sequence.html?docID:2004034888. The polynucleotide
XX of the invention are also useful in physical arrays of molecules and as
XX plant breeding markers. The recombinant DNA construct is useful for
XX improving plant tolerance to cold, heat, drought, herbicides, extreme
XX osmotic conditions, pathogens or pests, for manipulating growth rate in
XX plant cells by modification of the cell cycle pathway, for conferring

CC increased resistance to plant disease, for producing galat
CC lignin or plant growth regulators, for increasing the rat
CC recombination in plants, for improving yield by modificat
CC photosynthesis or carbohydrate, nitrogen or phosphorus us
CC or by providing improved plant growth and development und
CC stress condition or for modifying seed oil or protein yie
CC content. This is the amino acid sequence of a plant full
CC polypeptide that can be used in the recombinant DNA const
CC invention.
CC
CC SQ Sequence 173 AA;
Qy Query Match 40.8%; Score 401.5; DB 8; Length 1
Db Best Local Similarity 42.7%; Pred. No. 1.6e-32;
Matches 82; Conservative 28; Mismatches 45; Indels 4;
Qy 8 QETTSQAPEGPVMCKNLGCFSGSATMGLCSKYRETWQAMTALAK
Db 6 KEAGCQOPEGPILCINNCGFSGSAATMNMCSKCHKE-----MITKQED
Qy 67 -----TAAAVQPPAPVHETKLTCEVERTMIVPHQSSSYQODL
Db 60 VNGDGGKGPVLAASVNAVVP-----QVEQKTIY-----
Qy 116 SSIAAPSRPEPRRCGSKRVRGLTGPKRCGNYCALHRYSDKHTCTY
Db 101 AAVTPKAKEGPDRCACRKRVLGTGFSRCGNNYCSVHRYSBKHCQDF
Qy 176 ANPLVVAEKVVK 187
Db 161 ANPVVRAEKLDK 172
RESULT 11
ABM90052
ID ABM90052 standard; protein; 160 AA.
XX
XX AC ABM90052;
XX DT 02-JUN-2005 (first entry)
XX DE Rice abiotic stress responsive polypeptide SEQ ID NO:8389
XX KM abiotic stress tolerance; transgenic plant; plant; cereal
XX OS Oryza sativa.
XX OS WO2003008540-A2.
XX PN 30-JAN-2003.
XX PD 21-JUN-2002; 2002WO-US019668.
XX PF 22-JUN-2001; 2001US-0300112P.
XX PR 24-AUG-2001; 2001US-0314662P.
XX PR 26-SEP-2001; 2001US-0325277P.
XX PR 21-NOV-2001; 2001US-0332132P.
XX (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
XX Kreppe J, Briggs SP, Cooper B, Glazebrook J, Goff SA,
XX Moughamer T, Provart N, Rieke D, Zhu T;
XX WPI; 2003-248011/24.
XX
XX New stress-responsive nucleic acid, useful for altering t
XX responsiveness of a plant, e.g. cereal, to an abiotic str
XX stress, salt stress or osmotic stress.
XX
XX Claim 1, SEQ ID NO 8389; 89pp; English.
XX
XX The invention relates to novel abiotic stress responsive
XX and polypeptides. Also disclosed are vectors, expression

CC cells, and plants containing such polynucleotides. Also disclosed are
CC methods for using the polynucleotides and polypeptides to alter the
CC responsiveness of a plant to abiotic stress. The invention is useful in
CC agriculture. The nucleic acid is useful for determining whether a test
CC plant has been exposed to an abiotic stress condition. It is also useful
CC for selecting an agent that alters abiotic stress regulated
CC polynucleotide expression in a plant cell, and to identify a homolog or
CC ortholog to an abiotic stress responsive polynucleotide. The nucleic acid
CC molecule and the polypeptide encoded by it are useful in altering the
CC responsiveness of a plant to an abiotic stress, such as cold stress, salt
CC stress, osmotic stress or any of their combinations. The present sequence
CC is used in the exemplification of the invention

XX Sequence 160 AA;

QY Query Match 39.2%; Score 386; DB 7; Length 160;

Best Local Similarity 42.2%; Pred. No. 5.6e-31;
Matches 81; Conservative 23; Mismatches 50; Indels 38; Gaps 5;

DB 1 MATE---RVSOETTSQAPBGMCKNLGFFGSOATMGLCSKCYRETVMQAKMTALAQA 57
1 MAQESWKKEAETGVHTPEAPILCVNNCGFSGSMTMCKCYRDIV-KAK----- 51
QY 58 TQAAGATSATAAQQPAPVHETKLTCEVERTMIVPHOSSSYQODLVTPAAAPQAVKSS 117
52 -----TVAIVVEKKPLASLSSTPLVTB-----VTDDGSGSVADGKQ 87
DB 118 IAAAPREP--NRGSGCRKRVGLTGFKRCGNYCALHRYSDKTTCTDYKAAQGEALAK 175
88 VMEEDTPPKPSPNRCLSCSKRVKGLTGFKRCGCTGCSMHRVADSHKCTDYKQVREGQIAK 147
QY 176 ANPLVAEKKVK 187
148 QNPLVAKDKIRK 159
DB

RESULT 12
ABM87359
ID ABM87359 standard; protein; 169 AA.

XX AC ABM87359;
XX DT 02-JUN-2005 (first entry)
XX DE Rice abiotic stress responsive polypeptide SEQ ID NO:5605.
XX KW abiotic stress tolerance; transgenic plant; plant; cereal; agriculture.
XX OS Oryza sativa.
XX PN WO2003008540-A2.
XX PD 30-JAN-2003.
XX PF 21-JUN-2002; 2002WO-US019668.
XX PR 22-JUN-2001; 2001US-0300112P.
XX PR 24-AUG-2001; 2001US-0314662P.
XX PR 26-SEP-2001; 2001US-0325277P.
XX PR 21-NOV-2001; 2001US-0332132P.
XX PA (SYGN) SYNGENTA PARTICIPATIONS AG.
XX PI Krepis J, Briggs SP, Cooper B, Glazebrook J, Goff SA, Katagiri F,
XX PI Moughamer T, Provart N, Riecke D, Zhu T;
XX DR WPI; 2003-248011/24.
XX PT New stress-responsive nucleic acid, useful for altering the
XX PT responsiveness of a plant, e.g. cereal, to an abiotic stress such as cold
XX PT stress, salt stress or osmotic stress.
XX PS Claim 1; SEQ ID NO 5605; 89pp; English.

XX The invention relates to novel abiotic stress responsive p
CC and polypeptides. Also disclosed are vectors, expression c
CC cells, and plants containing such polynucleotides. Also di
CC methods for using the polynucleotides and polypeptides to
CC responsiveness of a plant to abiotic stress. The invention
CC agriculture. The nucleic acid is useful for determining w
CC plant has been exposed to an abiotic stress condition. It
CC for selecting an agent that alters abiotic stress regulate
CC polynucleotide expression in a plant cell, and to identify
CC ortholog to an abiotic stress responsive polynucleotide. T
CC molecule and the polypeptide encoded by it are useful in a
CC responsiveness of a plant to an abiotic stress, such as c
CC stress, osmotic stress or any of their combinations. The p
CC is used in the exemplification of the invention

XX Sequence 169 AA;

QY Query Match 39.2%; Score 386; DB 7; Length 160;

Best Local Similarity 42.2%; Pred. No. 6e-31;
Matches 81; Conservative 23; Mismatches 50; Indels

QY 1 MATE---RVSOETTSQAPBGMCKNLGFFGSOATMGLCSKCYRETVM 57
10 MAQESWKKEAETGVHTPEAPILCVNNCGFSGSMTMCKCYRDIV- 60
DB 58 TQAAGATSATAAQQPAPVHETKLTCEVERTMIVPHOSSSYQODLVTF 117
61 -----TVAIVVEKKPLASLSSTPLVTB-----VTDDGSGSVADGKQ 87
QY 118 IAAAPREP--NRGSGCRKRVGLTGFKRCGNYCALHRYSDKTTCTYD 175
97 VMEEDTPPKPSPNRCLSCSKRVKGLTGFKRCGCTGCSMHRVADSHKCTDY 156
QY 176 ANPLVAEKKVK 187
157 QNPLVAKDKIRK 168
DB

RESULT 13
AAG32536
ID AAG32536 standard; protein; 173 AA.

XX AC AAG32536;
XX DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 39269.
XX KW Protein identification; signal transduction pathway; metab
XX KW hybridisation assay; genetic mapping; gene expression cont
XX KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-00301439.
XX PR 25-FEB-1999; 99US-0121825P.
XX PR 05-MAR-1999; 99US-0123180P.
XX PR 09-MAR-1999; 99US-0123548P.
XX PR 23-MAR-1999; 99US-0125788P.
XX PR 25-MAR-1999; 99US-0126264P.
XX PR 29-MAR-1999; 99US-0126785P.
XX PR 01-APR-1999; 99US-0127462P.
XX PR 06-APR-1999; 99US-0128234P.
XX PR 08-APR-1999; 99US-0128714P.
XX PR 16-APR-1999; 99US-0129845P.
XX PR 19-APR-1999; 99US-0130077P.
XX PR 21-APR-1999; 99US-0130449P.

PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-013149P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0132863P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134419P.
PR 14-MAY-1999; 99US-0134421P.
PR 14-MAY-1999; 99US-013470P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135529P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-014091P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.

PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 28-JUL-1999; 99US-0145919P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 05-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 10-AUG-1999; 99US-0147935P.
PR 11-AUG-1999; 99US-0148171P.
PR 12-AUG-1999; 99US-0148319P.
PR 13-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 30-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157553P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158233P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0158293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.

PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 38.5%; Score 379; DB 3; Length 173;
Beet Local Similarity 43.5%; Pred. No. 3, 2e-30;
Matches 77; Conservative 34; Mismatches 46; Indels 20; Gaps 5;

QY 15 PEGPMCKKLCGFFSGQATMGCSCKYRETWQAKMTALAEQTAAQNTSTAAYPP 74
DB 12 PEGPKLTCTNNCGFFGSAATMNNCKCHKDKMLPQ-----QEGAKFASVSGTSSS----- 61
QY 75 AFWHETKLTCEVERIMIVPHOSSYQOQDLVTPAALAPQAVKSIAP--SRPE-PNRCG 130
DB 62 -----SNIKETFTALVDIETKSVEMTVSVQSSVQV-AEVVAPEBAAPKPKPSICT 115
QY 131 SCRKRVGELTGFCRCRCNTLYCALHRYSDKRTCTTYDKAQAQEAIAKANPLVAEKVVK 187
DB 116 TGNKKVGLTGFCRCRGSLEFCGTHRYADVHDCSFNTHAAQEAIAKANPLVNAEKLDK 172

RESULT 14

AAG49664 standard; protein; 173 AA.

AC AAG49664;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 62852.

XX Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
KM termination sequence.

OS Arabidopsis thaliana.

XX EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-00301439.

XX PR 25-FEB-1999; 99US-0121825P.

XX PR 05-MAR-1999; 99US-0123180P.

XX PR 09-MAR-1999; 99US-0123548P.

XX PR 23-MAR-1999; 99US-0125788P.

XX PR 25-MAR-1999; 99US-0126264P.

XX PR 01-APR-1999; 99US-0127462P.

XX PR 06-APR-1999; 99US-0128234P.

XX PR 08-APR-1999; 99US-0128714P.

XX PR 16-APR-1999; 99US-0129845P.

PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 04-MAY-1999; 99US-0132048P.
PR 05-MAY-1999; 99US-0132407P.
PR 06-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136029P.
PR 27-MAY-1999; 99US-0136332P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139452P.
PR 18-JUN-1999; 99US-0139452P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142920P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.

Accession	Score	Length	Identical	Mismatches	Indels
20-JUL-1999	99US-0144352P	173	100	0	0
20-JUL-1999	99US-0144632P	173	100	0	0
20-JUL-1999	99US-0144884P	173	100	0	0
21-JUL-1999	99US-0144814P	173	100	0	0
21-JUL-1999	99US-0145086P	173	100	0	0
21-JUL-1999	99US-0145088P	173	100	0	0
22-JUL-1999	99US-0145085P	173	100	0	0
22-JUL-1999	99US-0145087P	173	100	0	0
22-JUL-1999	99US-0145089P	173	100	0	0
22-JUL-1999	99US-0145192P	173	100	0	0
23-JUL-1999	99US-0145145P	173	100	0	0
23-JUL-1999	99US-0145218P	173	100	0	0
23-JUL-1999	99US-0145224P	173	100	0	0
26-JUL-1999	99US-0145216P	173	100	0	0
26-JUL-1999	99US-0145217P	173	100	0	0
27-JUL-1999	99US-0145913P	173	100	0	0
27-JUL-1999	99US-0145918P	173	100	0	0
28-JUL-1999	99US-0145951P	173	100	0	0
02-AUG-1999	99US-0146386P	173	100	0	0
02-AUG-1999	99US-0146388P	173	100	0	0
02-AUG-1999	99US-0146389P	173	100	0	0
03-AUG-1999	99US-0147038P	173	100	0	0
04-AUG-1999	99US-0147204P	173	100	0	0
05-AUG-1999	99US-0147302P	173	100	0	0
05-AUG-1999	99US-0147326P	173	100	0	0
06-AUG-1999	99US-0147303P	173	100	0	0
06-AUG-1999	99US-0147416P	173	100	0	0
09-AUG-1999	99US-0147493P	173	100	0	0
09-AUG-1999	99US-0147935P	173	100	0	0
10-AUG-1999	99US-0148171P	173	100	0	0
11-AUG-1999	99US-0148319P	173	100	0	0
12-AUG-1999	99US-0148341P	173	100	0	0
13-AUG-1999	99US-0148565P	173	100	0	0
13-AUG-1999	99US-0148684P	173	100	0	0
16-AUG-1999	99US-0149368P	173	100	0	0
17-AUG-1999	99US-0149175P	173	100	0	0
18-AUG-1999	99US-0149426P	173	100	0	0
20-AUG-1999	99US-0149722P	173	100	0	0
20-AUG-1999	99US-0149723P	173	100	0	0
20-AUG-1999	99US-0149929P	173	100	0	0
23-AUG-1999	99US-0149902P	173	100	0	0
23-AUG-1999	99US-0149930P	173	100	0	0
25-AUG-1999	99US-0150566P	173	100	0	0
26-AUG-1999	99US-0150884P	173	100	0	0
27-AUG-1999	99US-0151065P	173	100	0	0
27-AUG-1999	99US-0151066P	173	100	0	0
27-AUG-1999	99US-0151080P	173	100	0	0
30-AUG-1999	99US-0151303P	173	100	0	0
31-AUG-1999	99US-0151438P	173	100	0	0
01-SEP-1999	99US-0151930P	173	100	0	0
07-SEP-1999	99US-0152353P	173	100	0	0
10-SEP-1999	99US-0153070P	173	100	0	0
13-SEP-1999	99US-0153758P	173	100	0	0
15-SEP-1999	99US-0154018P	173	100	0	0
16-SEP-1999	99US-0154039P	173	100	0	0
20-SEP-1999	99US-0154799P	173	100	0	0

[illegible]

PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161932P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 38.5%; Score 379; DB 3; Length 173;
Best Local Similarity 43.5%; Pred. No. 3.2e-30;
Matches 77; Conservative 34; Mismatches 46; Indels 20; Gaps 5;

QY 15 PEGPVWCXKULCGPFGSOATMGKSCXCYRETVWQAKMTALAEQATOAQATSAATAVQPP 74
DB 12 PEGPKLCTNNGCFPGSAATNMKSKCHKDWLFO-----OEGAKFASAVSGTSSS---- 61
QY 75 APVHETKLTCEVERTMIVPHQSSSYQODLVTPAAAAPOAVKSSIAAP---SRPE-ENRCG 130
DB 62 ----SNIKETFTALVDIETKSVPEPMTVSVQSSVQVV-AEVVAPEEAKKXGSPRCT 115
QY 131 SCRKRVLGTGPKRCGKCNLYCALHRYSDKTCCTYDYKAGOEALAKANPLVNAEKVYK 187
DB 116 TCNKRVLGTGPKRCGSLFCGTHRYADVHDSCSFNYHAAOEALAKANPVVKAEXKLDK 172

Search completed: December 8, 2005, 16:29:10
Job time : 137 secs

This Page Blank (uspto)


```

A/Reference number: 225013
A/Accession: 149033
A/Status: Preliminary
A/Molecule type: DNA
A/Residues: 1-170 <PDB>
A/Cross-references: UNIPROT:Q9LXI5; UNIPARC:UPI000009DQFD; EMBL:AL353912; GSPDB:GN00061
A/Experimental source: cultivar Columbia; BAC clone F3C22
C/Genetics:
A/Gene: ATSP:F3C22.200
A/Map position: 3

Query Match      36.8%; Score 362.5; DB 2; Length 170;
Beet Local Similarity 37.8%; Fred. No. 2.9e-23;
Matches 73; Conservative 32; Mismatches 53; Indels 35; Gaps 4;

Oy 6 VSOETTSAPEGPWVCKYLCSFGSOATMGCSKCYRTVWQAKMTALAEATQAQATS 65
   ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 1 MAEHRRCOTPPSNRILCVNNCGFLSSATMNLCSNCTGDLCLK-----QQQSSSTKSTV 54

Oy 66 ATAAAVOPPAVEHETKLTCVEERTWIVPHOSSSYQOODLVTPAAAPQA-----VKSSIA 119
   ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 55 ESSLISVSPS-----SSSSSISPIIPLLKNPSVKLEVEPEKANIS 96

Oy 120 APS-----RPEPNRCGSCRKVGILTFGRCRGNLYCALHRSYSDKHTCTYYDKAQAQEAIA 174
   ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 97 LPTTEQNNQQQAPNRCTCRKRVGLTFGRKCRGTMFCVHRYPEIHGCSYDFPKSAGREIA 156

Oy 175 KANPLVVAEKVVK 187
   |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 157 KANPLVKAALK 169
   |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

```

```

RESULT 3
TS1098
hypochemical protein p85RF [imported] - Prunus armeniaca
C:Species: Prunus armeniaca
C:Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 09-Jul-2004
C:Accession: TS1098
R:Mdegule-A-Mdegule, D.; File-Lycaon, B.R.
A:Submitted to the EMBL Data Library, March 1999
A:Description: Molecular cloning and nucleotide sequence of a putative protein from apricot
A:Reference number: Z25296
A:Accession: TS1098
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-173 <MB>
A:Cross-references: A:PIROT:O9XH71; UNIPARC:UPI00000AC882; EMBL:AF139499; PIDN:AAD38146.
C:Genetics:
A:Note: p85RF

Query Match      36.7%; Score 361.5; DB 2; Length 173;
Beat Local Similarity 39.2%; Pred. No. 3.6e-23;
Matches 76; Conservative 26; Mismatches 53; Indels 39; Gaps 5;

Oy      8 QETTSQA-PEAPVWCKNLCGFFGSOATMGSCKYRETVMQAKMTALAEQ-----TQ 59
Db      4 EETGCGPPEBPIILCVNNCGFFGSVATRMNCSKCHKMDLKEBQAKLAASSFGINVGTS 63
Oy      60 AAQATSATAAIVOPPAVPHETKLTCVEERTWIVPHQSSYOODLVTPAAAPQ-----AV 114
Db      64 NSNGNEPVA--
Oy      115 KSIIPAAPSRA-PNRCGSCRKRVGLTGFKCRGNIYCALHRYSDKHTCTYDYKAAGEAI 173
Db      99 GSGSGEAKPEGPGRKRCGTCKNRVGLTGFCRCGHLCFAVHRHYSDBHDCPYDYHTHAARDVI 158
Oy      174 AKANPLVIAEKVK 187
Db      159 AKANPVVADKLEK 172

```

[illegible]

```

RESULT 5
C84779
hypothetical protein At2g36320 [imported] - Arabidopsis thalian
C|Species: Arabidopsis thaliana (mouse-ear cress)
C|Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change
C|Accession: C84779
R|Lin, X.; Kaul, S.; Rounsley, S.D.; Shee, T.P.; Benito, M.I.;
M.; Koo, H.; Moffitt, K.S.; Cronin, L.A.; Shen, M.; Vanhaken, S.E
eaus, D.; Nierman, W.C.; White, O.; Bisen, J.A.; Salzberg, S.L.
Nature 402, 761-768, 1999
A|Title: Sequence and analysis of chromosome 2 of the plant Ara
A|Reference number: A84420; MUID:20083487; PMID:10617197
A|Accession: C84779
A|Status: Preliminary
A|Molecule type: DNA
A|Residues: 1-161 <SNO>
A|Cross-references: UNIPROT:Q9SJM6; UNIPARC:UPI0000A045C; GB:A
A|Gene: At2g36320
A|Map position: 2

Query Match      36.2%; Score 356; DB 2; Length 161
Best Local Similarity 39.8%; Pred. No. 9.6e-23;
Matches 74; Conservative 29; Mismatches 53; Indels

6 VSOETTSQAPEGCPVWCNLCGFFSGQATMGCSKCYREYTMQAAKITALA 65
::: : : : : : : : : : : : : : : : : : : : : : : : :
1 MAEHRGCTPEGRHLCVNNCGFFSSATMNLCSNICYDDLCLIKQO----- 55
MAEHRGCTPEGRHLCVNNCGFFSSATMNLCSNICYDDLCLIKQO-----

66 ATAAAVOPPAVHETKLTCEVERMTVIPHOSSSYQODLVTPAAAPQ-- 122
S : : : : : : : : : : : : : : : : : : : : : : : : :
56 SLSTVIT--APVLE-----NYAALEIPTTKTBEK 94
S : : : : : : : : : : : : : : : : : : : : : : : : :

123 RPE-PNRGSGCRKRVGLTGFKRCGNLYCALHRRYSDRHTCTYDYKAAQ 181
- : : : : : : : : : : : : : : : : : : : : : : : : :
95 PPGAPNRCCTVGRKRVGLTGFRCRGCTTFCGSHRYPVHAGCTFDEKXAGR 154
- : : : : : : : : : : : : : : : : : : : : : : : : :

182 AEKVVK 187
- : : : : : : : : : : : : : : : : : : : : : : : : :
155 AAKLQK 160
- : : : : : : : : : : : : : : : : : : : : : : : : :

```


RESULT 6

D84674
hypothetical protein At2g27580 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:Accession: D84674

Rilin, X.; Kaul, S.; Rounleley, S.D.; Shea, T.P.; Bentlo, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.
neus, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: D84674

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-163 <STO>
A:Cross-references: UNIPROT:Q9ZNU9; UNIPARC:UPI00000AC021; GB:AE002093; NID:G3660274; PI
C:Genetics:
A:Gene: At2g27580
A:Map position: 2

Query Match 35.2%; Score 346.5; DB 2; Length 163;
Best Local Similarity 38.2%; Pred. No. 6e-22;
Matches 76; Conservative 30; Mismatches 44; Indels 49; Gaps 6;

QY 1 MATERNOSQETTSQAPBGPVCKNLGFGSSQATMGLCSKCYRETYMQAKMTALAEOA-TQ 59
DB 1 MAEEHRLQD-----PRLCANNGCFPGSTATQNLCSKCPDLQHQQNSSTAKHALTQ 52
QY 60 AAGATSAT-AAVOPAPVHEHTKTCCEVERMTIVPHQSSYQODLVTPAAAPQAV--KS 116
DB 53 SLAAAGAAASSSVSPPP-----PPADSKETVEAKS 83
QY 117 STAAFSRPE-----PVRCSCKRKGVLGTGFKCRGNYLCALHRYSDKHTCTDYKAA 168
DB 84 EGRBAAPPEAAGPPDPGRCLTCRRRVGITGFRCKGVPFGCTHRYAAGHCSPDFKAM 143
QY 169 GQBAIAKANPLVVAEKVK 187
DB 144 GKDIAKANPIVKADELK 162

RESULT 7

T04577
hypothetical protein T12H17.210 - Arabidopsis thaliana
N:Alternate names: hypothetical protein F7H19.10
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C:Accession: T04577; T05114

R:Bevan, M.; Hilbert, H.; Brun, M.; Holzer, E.; Brandt, A.; Duesterhoeft, A.; Bancroft,
submitted to the Protein Sequence Database, February 1998
A:Reference number: Z15377
A:Accession: T04577

A:Molecule type: DNA
A:Residues: 1-176 <BEV>
A:Cross-references: UNIPROT:Q49663; UNIPARC:UPI00000ABF74; EMBL:AL021635
A:Experimental source: cultivar Columbia; BAC clone T12H17
R:Bevan, M.; Peters, S.A.; van Staveren, M.; Dirkse, W.; Stiekema, W.; Bancroft, I.; Men
submitted to the Protein Sequence Database, July 1998
A:Reference number: Z15399
A:Accession: T05114

A:Molecule type: DNA
A:Residues: 1-176 <BEW>
A:Cross-references: UNIPARC:UPI00000ABF74; EMBL:AL031018
A:Experimental source: cultivar Columbia; BAC clone F7H19
C:Genetics:
A:Map position: 4
A:Note: T12H17.210; F7H19.10

Query Match 34.3%; Score 338; DB 2; Length 176;
Best Local Similarity 41.2%; Pred. No. 3.3e-21;

Matches 75; Conservative 29; Mismatches 58; Indels 6;

QY 10 TTSSQAPBGPVCKNLGFGSSQATMGLCSKCYRETYMQAKMTALAEOA 69
DB 12 TOSQASE-PKLCVKGCGFGSPSNDLCSKCYRGICAEBAOTAVAKAA 70
QY 70 --AVOPAPVHEHTKTCCEVERMTIVPHQSSYQODLVTPAAAPQAVK 126
DB 71 LFIAPPPAVVPEK--PEAAVVVSAEBS-----SAVPEANE 114
QY 127 NRCGSCRRKRVGLTGFCKRCRGNLYCALHRYSDKHTCTDYKAAQOBAIA 186
DB 115 NRCLCNKKVGMGRKCKGSPFCGEHRYPERHDCSPDKVGRGEIAH 174
QY 187 KF 188
DB 175 RF 176

RESULT 8

T21254
hypothetical protein F22D6.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change
C:Accession: T21254

R:Wilkinson, J.
submitted to the EMBL Data Library, April 1996
A:Reference number: Z19397
A:Accession: T21254
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA

A:Residues: 1-189 <WIL>
A:Cross-references: UNIPROT:Q19723; UNIPARC:UPI000007FF68; EMBL
A:Experimental source: clone F22D6
C:Genetics:
A:Gene: CESP:F22D6.2
A:Map position: 1
A:Intons: 51/1, 106/2, 141/1

Query Match 31.8%; Score 313; DB 2; Length 18;
Best Local Similarity 35.4%; Pred. No. 4.3e-19;
Matches 68; Conservative 31; Mismatches 69; Indels 6;

QY 8 QETTSQAPBGPVCKNLGFGSSQATMGLCSKCYRETYMQAKMTALAEOA 67
DB 5 QOQOACTAPS-----CRAGCGFGASATBEGYSCQCFKNTILKROODTVRLT 60
QY 68 AAAYQ-PPAPV-----HETKLTCEVERMTIVPHQSSYQOD 113
DB 61 SSALSKSPSSVDMCKKAIVSVSDETAKNDCE---DIIIVCDQINDO: 115
QY 114 VKSSIAAPSRPEPNRCGSCRRKRVGLTGFCKRCRGNLYCALHRYSDKHTCT 173
DB 116 ITVDVPPVVK--KARNCMKRKRVLGTGSCRCGGIYCGDHHYDDQAHNC 174
QY 174 AKANPLVVAEKV 185
DB 175 KKNPPVVVSDVKV 186

RESULT 9

T11846
pathogenesis-related protein 3 - kidney bean
C:Species: Phaseolus vulgaris (kidney bean)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change
C:Accession: T11846

R:Sharma, Y.K.; Hinojos, C.M.; Mehdy, M.
Mol. Plant Microbe Interact. 5, 89-95, 1992
A:Title: CDNA cloning, structure and expression of a novel pat
A:Reference number: Z17354; MUID:92288407; PMID:1600239
A:Accession: T11846
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA

C/Species: Bos primigenius taurus (cattle)
C/Date: 04-Sep-1997 #sequence_revision 04-Sep-1997 #text_change 09-Jul-2004
C/Accession: I47035
R/Shima, D.T.; Saunders, K.B.; Gougos, A.; D'Amore, P.A.
Differentiation 58, 217-226, 1995
A/Title: Alterations in gene expression associated with changes in the state of endothel
A/Reference number: I47034; MID:95228954; PMID:7713329
A/Accession: I47035
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-46 <SH1>
A/Cross-references: UNIPROT:Q28900; UNIPARC:UPI000008679C; GB:S77733; NID:g998679; PIDN:

Query Match 18.0%; Score 177; DB 2; Length 46;
Best Local Similarity 69.8%; Pred. No. 2.4e-08;
Matches 30; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

Qy 143 CRCGNLYCALHRYSDKHTCTYDYKAGQEAIAKANPLVVAEKV 185
Db 1 CRCGNLFCGLHRYSDKHTCTYDYKAGQEAIAKANPLVVAEKI 43

RESULT 14

T18125
hypothetical protein A623L - Chlorella virus PBCV-1
C/Species: Chlorella virus PBCV-1
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C/Accession: T18125
R/Graves, M.V.; Van Etten, J.L.
submitted to the EMBL Data Library, May 1999
A/Reference number: Z18806
A/Accession: T18125
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-67 <GRA>
A/Cross-references: UNIPROT:Q41105; UNIPARC:UPI000006F20; EMBL:U42580; NID:g4028896; PI
A/Experimental source: specific host Chlorella strain NC64A
C/Genetics:
A/Note: A623L

Query Match 12.6%; Score 124; DB 2; Length 67;
Best Local Similarity 36.7%; Pred. No. 0.00092;
Matches 22; Conservative 10; Mismatches 18; Indels 10; Gaps 1;

Qy 127 NRCGCRKRVGLTGFRCGNYCALHRYSDKHTCTYDYKAGQEAIAKANPLVVAEKV 186
Db 10 SRCEICRKKTKGLGFCVCKGHTFCERGRIMESHSC-----PTLQAKERIILEKV 59

RESULT 15

S43586
F26F3.4 protein (clone F26F3) - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 07-Sep-1994 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C/Accession: S43586
R/Suleton, J.
submitted to the EMBL Data Library, April 1994
A/Reference number: S43582
A/Accession: S43586
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-341 <SUL>
A/Cross-references: UNIPROT:Q20899; UNIPARC:UPI000008295B; EMBL:Z32681; NID:g474790; PID
C/Genetics:
A/introns: 33/3; 101/2; 192/2; 232/3; 277/3

Query Match 11.4%; Score 112.5; DB 2; Length 341;
Best Local Similarity 28.6%; Pred. No. 0.041;
Matches 24; Conservative 12; Mismatches 33; Indels 15; Gaps 2;

Qy 112 QAVKSIAPSRPE-----PVRCSGRKRVGLT--GFKRCGNLYCALHRYSDKHTCTYDYKAGQEAIAKANPLVVAEKV 186
Db 246 QCRKSNVLAFFSBEELLENLKKFEARKTKCNTCFKKLSAAGOTMCKCLRIFCDRHRHP 305

Qy 157 DKHTCTYDYKAGQEAIAKANPLV 180
Db 306 KNHTCTYDYKQDGRNRKLIKKNNSKV 329

Search completed: December 8, 2005, 16:32:58
JOD time : 48 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 2005, 12:58:00 / Search time 159 Seconds

(without alignments)
834.210 Million cell updates/sec

Title: US-10-716-089-20

Sequence: 1 MATEVSQETTSQAREGPVM.....GGEALAKANPLVAAEKVKVF 188

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot 05.80:*

1: uniprot_sprot:*

2: uniprot_tr embl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	419.5	42.6	171	2	Q69WY4_ORYSA
2	416.5	42.3	171	2	Q6H7P8_ORYSA
3	415.5	42.2	171	2	Q6VAG2_ORYSA
4	413	42.0	233	2	Q6R318_MALZE
5	386	39.2	160	2	Q852K5_ORYSA
6	384.5	39.1	161	2	Q7Y1W9_ORYSA
7	384	39.0	173	2	Q8H0X0_ARATH
8	379	38.5	173	2	Q9SYC3_ARATH
9	369	37.5	211	2	Q6E1P3_XENLA
10	365	37.1	211	2	Q6DF90_XENLA
11	365	37.1	213	1	Z20D2_MOUSE
12	364	37.0	213	2	Q6PBR4_BRARE
13	364	37.0	213	2	Q6PBR4_BRARE
14	363.5	36.9	212	2	Q5ZUR4_CHICK
15	362.5	36.8	170	2	Q9LX15_ARATH
16	361.5	36.7	173	2	Q9LX15_ARATH
17	361.5	36.7	175	1	ZF2N1_ARATH
18	361.5	36.7	212	2	Q7ZU97_BRARE
19	361	36.7	224	2	Q6Z541_ORYSA
20	356	36.2	161	1	ZF2N2_ARATH
21	352	35.8	169	2	Q94B40_ARATH
22	352	35.8	223	2	Q6DGF4_RAT
23	352	35.8	223	2	Q9DCH6_MOUSE
24	350.5	35.6	210	2	Q4G2S7_PORCU
25	347	35.3	611	2	Q4SD48_TERNG
26	346.5	35.2	163	2	Q9ZNU9_ARATH
27	346	35.2	207	2	Q4TU6_TERNG
28	344.5	35.0	224	2	Q4T017_TERNG
29	343.5	34.9	169	2	Q852K6_ORYSA
30	343.5	34.9	208	2	Q9GZY3_HUMAN
31	343	34.9	170	2	Q84PD8_ORYSA

32	343	34.9	223	2	Q9ER79_MOUSE
33	341.5	34.7	208	2	Q5R7S6_PONPY
34	340	34.6	199	2	Q9VHF4_DROME
35	338.5	34.4	148	2	Q942F8_ORYSA
36	338	34.3	168	1	ZF2N1_ARATH
37	338	34.3	176	2	Q49663_ARATH
38	338	34.3	206	2	Q8WQ62_DROME
39	337.5	34.3	164	1	ISAP1_ORYSA
40	337.5	34.3	202	2	Q96038_CIOGA
41	335.5	34.1	198	2	Q7QAK6_XENLA
42	334	33.9	201	2	Q6P840_XENTR
43	333.5	33.9	154	2	Q6H754_ORYSA
44	332.5	33.8	208	2	Q9BQF7_HUMAN
45	327.5	33.3	208	2	Q6T1P0_HUMAN

ALIGNMENTS

RESULT 1	Q69WY4_ORYSA	PRELIMINARY;	PRT;	171 AA.
AC	Q69WY4;			
DT	25-OCT-2004 (TrEMBLrel. 28, Created)			
DT	25-OCT-2004 (TrEMBLrel. 28, Last sequence update)			
DT	01-FEB-2005 (TrEMBLrel. 29, Last annotation update)			
DE	Putative multiple stress-responsive zinc-finger protein.			
GN	Name=P0417G12.11; Synonyms=P0429G06.28;			
OS	Oryza sativa (japonica cultivar-group).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Trac			
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae			
OC	Enarthroideae; Oryzeae; Oryza.			
OX	NCBI_TaxID=39947;			
NP	NUCLEOTIDE SEQUENCE.			
RA	Sasaki T., Matsunoto T., Yamamoto K.;			
RT	"Oryza sativa nipponbare (GA3) genomic DNA, chromosome 6, 1			
RT	clone:P0417G12.11;			
RL	Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.			
RP	NUCLEOTIDE SEQUENCE.			
RA	Sasaki T., Matsunoto T., Yamamoto K.;			
RT	"Oryza sativa nipponbare (GA3) genomic DNA, chromosome 6, 1			
RT	clone:P0429G06.28;			
RL	Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AP003711; BAD3553.1; -; Genomic DNA.			
DR	EMBL; AP003711; BAD3553.1; -; Genomic DNA.			
DR	Gramene; 069WY4; -;			
DR	GO; GO:0003677; F:DNA binding; IEA.			
DR	GO; GO:0008270; F:zinc ion binding; IEA.			
DR	InterPro; IPR002653; Znf_A20.			
DR	InterPro; IPR000058; Znf_A20.			
DR	Pfam; PF01754; zf-A20; 1.			
DR	Pfam; PF01428; zf-AN1; 1.			
DR	SMART; SM00259; Znf_A20; 1.			
DR	SMART; SM00154; Znf_AN1; 1.			
KW	Zinc-finger.			
SO	SEQUENCE 171 AA; 18402 MW; 5E425E95741816C4 CRC64;			
Query Match	42.6%; Score 419.5; DB 2; Length			
Best Local Similarity	45.6%; Pred. No. 7.4e-27;			
Matches	82; Conservative 26; Mismatches 59; Indels			
QY	8 QETTSQAREGVMCKNLGFFGSGATMGLCSKCYRETYWQAKMTALAE			67
DB	4 KETGCGQEGEGILICINMGFGFGSAATNMCKCHKEMTMKQEQAKLA-			61
QY	68 AAAYQPPAPVHEITKLTCEVERTMTVPHQSSSTQDDLTTPAAAPQAVK			127
DB	62 DSGKEPPIIAGAAVAVAQVEVKTLV-----AQPATIAQSGSEG			110
QY	128 RCGSCRKRVGLTGFCRCGNLYCALHRSYSDKHTCTYDYKAAQGEALIAV			187

Db 111 RCTCKRVRGLTGFNRCNGNLXCAWHRRSDKHDCQFDVRYTARDAIKANPVYAEGLDK 170

RESULT 2

ID	Q6H7P8_ORYSA	PRT; 173 AA.
AC	Q6H7P8_ORYSA PRELIMINARY;	
AC	Q6H7P8;	
DT	05-JUL-2004 (TrEMBLrel. 27, Created)	
DT	05-JUL-2004 (TrEMBLrel. 27, Last sequence update)	
DT	05-JUL-2004 (TrEMBLrel. 27, Last annotation update)	
DE	Putative zinc-finger protein.	
GN	Name=OJ1225_F07.15;	
OS	Oryza sativa (japanese cultivar-group).	
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;	
OC	Ehrhartoideae; Oryzeae; Oryza.	
OX	NCBI_taxid=39947;	
RN	[1]	
RP	NUCLEOTIDE SEQUENCE.	
RA	Sasaki T., Matsumoto T., Yamamoto K.;	
RL	Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.	
DR	EMBL: AP004184; BAD25251.1; -; Genomic_DNA.	
DR	Gramene; Q6H7P8; -	
DR	GO; GO:0003677; P:DNA binding; IEA.	
DR	GO; GO:0008270; F:zinc ion binding; IEA.	
DR	InterPro; IPR002653; ZnF_A20.	
DR	InterPro; IPR000058; ZnF_AN1.	
DR	Pfam; PF01754; zf-A20; 1_	
DR	Pfam; PF01428; zf-AN1; 1.	
DR	SMART; SM00259; ZnF_A20; 1.	
DR	SMART; SM00154; ZnF_AN1; 1.	
DR	zinc-finger.	
Q	SEQUENCE 173 AA; 18540 MW; 5B3BC00752359ADE CRC64;	

Query Match	Similarity	42.3%	Score 416.5	DB 2	Length 173
Beat Local	Similarity	42.7%	Pred. No. 1.3e-26		
Matches	82	Conservative	28	Mismatches	47
				Indels	35
				Gaps	3

Query	8	QETTSOAPBEGPVMCKNLGCFPGSQATMGLCSKCYRETYMOAKMTALAEQATQ-----	60
Db	4	KEACCCQPEGIIILCINNGPFGSAAITNMGSKCHEMIMKEBQALTAASISDSINGCDG	63
Qy	61	-----AQTSAITAAVOPPAVHETKLTCEVERIMIIPHOSSSYOQDILVTPAAAAPQAVK	115
Db	64	GKEHIVAAAGSTAAVAV-----AAVEAKTLV-----VQPTDVAGTSEE	100
Qy	116	SSIAAPSRPEENRGSSCKRKGTLGFKRCGNLYALHRYSDKHTCTYDYAAAGCEATAK	175
Db	101	VAVPVKVEGNGRCATCKRKRVGLTGFGNRCGNMYCALHRYSDKHCEQDFRYTAARDATAK	160
Qy	176	ANPLVVAEKVVK	187
Db	161	ANPVVKAEEKDK	172

RESULT 3	Q6VAG2_ORYZA	Q6VAG2_ORYSA	PRELIMINARY	PRT	171 AA.
AC	Q6VAG2_				
DT	05-JUL-2004	(TREMBLrel. 27, Created)			
DT	05-JUL-2004	(TREMBLrel. 27, Last sequence update)			
DT	05-JUL-2004	(TREMBLrel. 27, Last annotation update)			
OC		Zinc-finger protein.			
OS		Oryza sativa (indica cultivar-group).			
OC		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC		Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;			
OC		Ehrhartoideae; Oryzeae; Oryza.			
NCBI		NCBI_TaxID=39946;			
OX		[1]			
RN					
RP					
RC					
EC					
RA					
FL					
Submitted					

Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.

DR EMBL AY345559; AAC08434.1; -, mRNA.
DR Gmeme; Q6VAG2; F:DNA binding; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR InterPro; IPRO02653; ZnF_A20.
DR InterPro; IPRO0058; ZnF_AN1.
DR Pfam; PF01754; zf-A20_1_AN1.
DR Pfam; PF01428; zf-AN1_1.
DR SMART; SM00259; ZnF_A20; 1.
DR SMART; SM00154; ZnF_AN1; 1.
KW Zinc-finger.
SEQUENCE 171 AA; 18401 MW; 07760A1F54108448 CRC64;

	Query Match	42.2%	Score 415.5;	DB 2;	Length :	
	Beeb Local Similarity	45.0%	Pred. No. 1.6e-26;			
	Matches	81;	Conservative	27;	Mismatches	59;
					Indels	2
Qy	8 QETTSOAPBEGPVMCKNLGFFGSOATNGLCSKCYRETFYMOANMTALAC					67
	: : : : : : : : : : : : : : : : : : : : :					
Dd	4 KETCCQQPKGPILCINCGFFPGSAATNMCKCHKEMTKBOAKLA--					61
	:					
Qy	68 AAAVQPAPVHETLTCEVERTMTVPHQSSSYOODLVTPAALAPOAVKS					127
	: : : : : : : : : : : : : : : : : : : : :					
Dd	62 DSGKEPIIAGHAFAVAQAQVEVKTLV-----AQPALAGSBEV					110
	:					
Qy	128 RCGSCRKRVLGTGFCKRCGNLYCALHRYSDMHTTLYDYKAAGEAIANA					187
	: : : : : : : : : : : : : : : : : : : : :					
Dd	111 RCSTCRKRVLGTGNCRCGNLYCAMHRSYDGDGOCFDVTRARAPIAAD					170
	: : : : : : : : : : : : : : : : : : : : :					

```

RESULT 4
Q6R3I8_MAIZE
ID Q6R3I8_MAIZE PRELIMINARY; PRT; 233 AA

```

DT 05-JUL-2004 (TEMBRE). 27, (Created)
 DT 05-JUL-2004 (TEMBRE). 27, Last sequence update)
 DT 05-JUL-2004 (TEMBRE). 27, Last annotation update)
 DE Zea mays (Maize).
 OS Zea mays (Maize).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tric
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae.
 OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
 OX NCBI_TaxId=4577;
 OX [1]
 RP NUCLEOTIDE SEQUENCE.
 RP Zhang Z.-X., Tang W.-H., Tao Y.-S., Zheng Y.-L.;
 RA Submitted (JAN-2004) to the EMBL/Genbank/DBJ databases.
 RL EMBL; AY515607; AAS00453.1; -; mRNA.
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0008270; F:zinc ion binding; IEA.
 DR InterPro; IPR000653; ZnF_A20.
 DR InterPro; IPR000058; ZnF_ANL.
 DR Pfam; PF01754; Zf-A20; 1.
 DR Pfam; PF01428; ZF-ANL; 1.
 DR SMART; SM00259; ZnF_A20; 1.
 DR SMART; SM00154; ZnF_ANL; 1.
 SQ SEQUENCE 233 AA; 25019 MW; 5561DAC11C1478C6 CRC64;

Query Match	Best Local Similarity	42.0%;	Score 413;	DB 2;	Length 233
Matches	81; Conservative	30;	Pident	No.3.5e-26;	Indels 58;
QY	8 QETTSQAPGEGVMCKNLGFFGSOATMGLCSCKCYRETYMAQMTALAEQ				67
DB	63 KEACGQAPGEGPILCINMGFFGSAATMMCKCKCHEMTTKDQAKLA--				120
QY	68 AAAYQAPAPVHETKLTCEVERMTLVPHQSSSYQODLVTPA-AAAPQAVK				126
DB	121 DAVMEPGS-----AGSNTVVAVAQVELQKKXVQAPDAVAEENEGV				171
QY	127 NRCSCKRIVGLTGPKRCRGNTLYCALHRYSPDKHTCTYYKAAQGEALAK				186
DB	172 NRCACRRIVGLTGPNCGNTLYCALHRYSPDKHCKEYRTPAARALAK				231

OY 187 K 187
DB 232 K 232

RESULT 5

O852K5_ORYSA PRELIMINARY; PRT; 160 AA.

AC O852K5;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-FEB-2005 (TREMBlrel. 29, Last annotation update)
DE Putative zinc finger protein (Putative zinc finger transcription factor ZFP38)
GN Name=OSUNB0060J21.18; Synonym=ZFP38;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
NCBI_TaxId=39947;

NUCLEOTIDE SEQUENCE.

RA Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Jones K.M.,
RA Overton II L.L., Telford T., Kim M.M., Bera J.J., Jin S.S.,
RA Fadrosch D.W., Tallon L.J., Koo H., Ziemann V., Hsiao J., Blunt S.,
RA Vanaken S.S., Riedmiller S.B., Uterback T.T., Feldblum T.V.,
RA Yang Q.Q., Haas B.J., Suh B.B., Peterson J.J., Quackenbush J.,
RA White O., Salzberg S.L., Fraser C.M.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
[2]
RN NUCLEOTIDE SEQUENCE.

RA Buell R.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.

NUCLEOTIDE SEQUENCE.

RA Huang J., Zhang H.-S.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC090871; AA037974.1; -; Genomic_DNA.
DR EMBL; AY377427; AA083587.1; -; mRNA.
DR Gramene; O852K5; -;
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR InterPro; IPR002653; Znf_A20.
DR InterPro; IPR000058; Znf_AN1.
DR Pfam; PF01428; Zf-AN1; 1.
DR Pfam; PF01428; Zf-AN1; 1.
SQ SEQUENCE 160 AA; 17532 MW; 1A10D44CF546B232 CRC64;

Query Match 39.2%; Score 386; DB 2; Length 160;

Best Local Similarity 42.2%; Pred. No. 4e-24;

Matches 81; Conservative 23; Mismatches 50; Indels 38; Gaps 5;

OY 1 MATE---RVSOETTQAPGPGVPMCKNLGCFGSGQATMGLCKCYRETVMQAATLAAQA 57
DB 1 MAOESWKKEAEETGHTPEAPILCVNCGFGSSMTNNKSCYRDFV-KAK----- 51
OY 58 TQAAGTATAAAVQPPAPVHETKLTCEVERTMIVPHOSSYQODLVTPAALPAQAVKS 117
DB 52 -----TVATVVEKKPLASLSTPLVTE-----VTDGSGSVADKQ 87
OY 118 IAAPEPPEP--NRGSCSKRKGVLTFKRCRCGNLYCALHRYSPKATCTYDYRAAGCEAIK 175
DB 88 VVEEDTPEPSPRCLSCSKKRGVLTGFKRCGCTFGCSMRVADSHKCTPDYKQVGREQIAK 147
OY 176 ANPLVVAEKVKV 187
DB 148 QNPLVKAADKITK 159

RESULT 6
O7Y1W9_ORYSA PRELIMINARY; PRT; 161 AA.

AC O7Y1W9;
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT 01-FEB-2005 (TREMBlrel. 29, Last annotation update)
DE Putative zinc finger transcription factor ZFP33 (Putative protein 216).
GN Name=OSUNB0050P10.11;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Trac;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae
OC Ehrhartoideae; Oryzaceae; Oryza.
NCBI_TaxId=39947;

NUCLEOTIDE SEQUENCE.

RA TISSUE=Root;
RC Huang J., Zhang H.-S.;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
[2]
RN NUCLEOTIDE SEQUENCE.

RA Sasaki T., Matsunoto T., Katayose Y.;
RT "Oryza sativa nippondate (GA3) genomic DNA, chromosome 7, F
RT clone:OSUNB0050P10.";
RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY282740; AAP37480.1; -; mRNA.
DR EMBL; AP005840; BAD11780.1; -; Genomic_DNA.
DR Gramene; O7Y1W9; -;
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR InterPro; IPR002653; Znf_A20.
DR InterPro; IPR000058; Znf_AN1.
DR Pfam; PF01754; Zf-A20; 1.
DR Pfam; PF01428; Zf-AN1; 1.
SQ SEQUENCE 161 AA; 17680 MW; 1443570BF15DD149 CRC64;

Query Match 39.1%; Score 384.5; DB 2; Length

Best Local Similarity 42.4%; Pred. No. 5.4e-24;

Matches 81; Conservative 23; Mismatches 52; Indels 5;

OY 1 MATE---RVSOETTQAPGPGVPMCKNLGCFGSGQATMGLCKCYRETVM 58
DB 1 MAOESWKKEAEETGHTPEAPILCVNCGFGSSMTNNKSCYRDFV- 53
OY 59 QAAGTATAAAVQPPAPVHETKLTCEVERTMIVPHOSSYQODLVTP? 118
DB 54 -----APV-----VEKKAFTPASSKTPLEPAKI? 89
OY 119 AA--PSRPEPNRCSCKRKGVLTFKRCRCGNLYCALHRYSPKATCTYD 176
DB 90 AAQEPKPRPSNRCLSCSKKRGVLTGFKRCGCTFGCSMRYTANHCTFD 149
OY 177 NPLVVAEKVKV 187
DB 150 NPVVAEKVK 160

RESULT 7

O8H0X0_ARATH PRELIMINARY; PRT; 173 AA.

AC O8H0X0;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Expressed protein.
GN Name=At1G51200;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Trac;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicot
OC eustosida II; Brassicales; Brassicaceae; Arabidopsids.
OC NCBI_TaxId=3702;

NUCLEOTIDE SEQUENCE.

RA Nguyen M., Karlin-Neumann G., Southwick A., Tripp M., Mir,
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Che.

RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Kidney;
 RA Klein S., Gerard D.S.;
 RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC081266; AAH81266.1; -, mRNA.
 DR SMR; 066193; 138-194.
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0008270; F:zinc ion binding; IEA.
 DR InterPro; IPR002653; ZnF_A20.
 DR InterPro; IPR000058; ZnF_AN1.
 DR Pfam; PF01754; zf-A20; 1.
 DR Pfam; PF01428; zf-AN1; 1.
 DR SMART; SM00259; ZnF_A20; 1.
 DR SMART; SM00154; ZnF_AN1; 1.
 SQ SEQUENCE 211 AA; 22887 MW; F08CBF1EC3788964 CRC64;
 Query Match 37.5%; Score 369; DB 2; Length 211;
 Best Local Similarity 37.9%; Pred. No. 1,4e-22;
 Matched 81; Conservative 29; Mismatched 64; Indels 40; Gaps 7
 QY 6 VSOETTSQAPESPVWCKNLGCFPGSQATGTCSCKCYRRTVMQKMTALAEQ-ATQAQAT 64
 DB 1 MAQG-TNQTP-GPMLCNTGCCGFGYGNPRITNGKMSVCYKHELRQNSGRISPMKAAGSNP 58
 QY 65 SATAAAVQPPAPVHEHTKLTCE-----VERTMTVPHQ 96
 DB 59 SAESAASVQRV---ETSLNCEGAVGGLSDKSHNTPLAALPYTQMTMSISREHVASPK 114
 QY 97 SSYQDDLT---PAAAPQAVKSSLAASRREP--NRGSCRKRVGLTFKRCGNLYCA 151
 DB 115 TETSEPVVTPSPSPVSQSPSTSLDEKAELELPFKKQNRCEFMCRKXIGLTFDRCNLTFCG 174
 QY 152 LHRYSDKTKTCYDYKACAGQEMATKANKPVAEKV 185
 DB 175 LHRYSDKNCPDYKAEAAKIRKENPVVAEKI 208
 RESULT 10
 06DF90_XENLA PRELIMINARY; PRT: 211 AA.
 ID 06DF90_XENLA PRELIMINARY;
 AC Q6DF90_2
 DT 25-OCT-2004 (TREMBLrel. 28, Created)
 DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
 DE Zs20d2-prov proteoin.
 GN Name=zs20d2-prov;
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipridae;
 OC Xenopodina; Xenopus; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Oocyte;
 RX MEDLINE=23388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.D., Feingold E.A., Grouse L.H., Berger J.G.,
 RA Klapper R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
 RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Datchenko L., Marzella K., Farmer A.A., Rubin G.M., Hong L.,
 RA StACPleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udell T.B., Toshiyuki S., Carlinci P., Zhang C.,
 RA Raba S.S., Logguelano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McEwen P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting R.W., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butcherfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
 RA Schnerf A., Schein J.E., Jones S.J.M., Marra M.A.;

RT	"Generation and initial analysis of more than 15,000 full
RT	and mouse cDNA sequences."
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN	[2]
RP	NUCLEOTIDE SEQUENCE.
RC	TISUE=Oocytes;
RX	MEDLINE=223341133; PubMed=12454917; DOI=10.1002/dvdy.10174
RA	Klein S.L., Strussberg R.L., Wagner L., Pontius J., Clifton
RA	Richardson P.;
RT	"Genetic and genomic tools for Xenopus research: The NIH
RT	initiative."
RL	Dev. Dyn. 225:384-391(2002).
RN	[3]
RP	NUCLEOTIDE SEQUENCE.
RC	TISUE=Oocytes;
RA	Klein S., Gerhard D.S.;
RL	Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR	EMBL; BC076851; AAT76851.1; -, mRNA.
DR	SMR; Q6DF90; 138-194.
DR	GO; GO:0003677; F-DNA binding; IEA.
DR	GO; GO:0008270; F:zinc ion binding; IEA.
DR	InterPro; IPR002653; Znf A20.
DR	InterPro; IPR000058; Znf_AN1.
DR	Pfam; PF01754; ZF-A20; 1.
DR	Pfam; PF01428; ZF-AN1; 1.
DR	SMART; SM00259; ZNF A20; 1.
DR	SMART; SM00154; ZNF AN1; 1.
SO	SEQUENCE 211 AA; 22974 MW; A6FC08103F36BD1D CRC64;
Query Match	37.1%; Score 365; DB 2; Length 211
Best Local Similarity	36.9%; Pred. No.2.9e-22;
Matches	79; Conservative 32; Mismatches 63; Indels
Qy	6 VSOETTSQAEGEVGMCKNTCGFSGQATMGTCSCYERYMGAOKMTAL
Db	1 MAQG-INQTP-GEMLCTCTGCGFYGNPRTMGCVCKEHLQRONSGRIT
Qy	65 SATAAAOQPAPAPHETLTCTCE-----VERTWIVPQK
Db	59 STSASAIQRV-----ETSLNCEGAAGALSDKSNTPIASLPVYQGMEM
Qy	108 AAAPQAV---KSIAPPS-----RPEPRCGSCRKRGTLT
Db	115 TETSEPVPTVPSPSVAPQSTSLNEBKEPELPPKPKKKRCPEWCKKIGLT
Qy	152 LHRYSDDHGTCTTYKYKAQGAIAIKAPLVAAYKV 185
Db	175 LHRYSDDHGNCPYDYKAAAKAKIRKENPVVAEXT 208
RESULT 11	
Z2D02_MOUSE	STANDARD; PRT; 213 AA.
AD	Z2D02_MOUSE
IC	O88878;
DT	28-FEB-2003 (Rel. 41, Created)
DT	28-FEB-2003 (Rel. 41, Last sequence update)
DT	10-MAY-2005 (Rel. 47, Last annotation update)
DE	Zinc finger A20 domain containing protein 2 (Zinc finger
GN	Name=Za20d2; Synonyms=Zfp216; Znf216;
OS	Mus musculus (mouse).
OC	Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Eutele
OC	Mammalia; Eutheria; Eumarchontoglires; Glires; Rodentia; S
OC	Muroidea; Muridae; Murinae; Mus.
OX	NCBI_TaxId=10090;
RN	[1]
RP	NUCLEOTIDE SEQUENCE, AND TISSUE SPECIFICITY.
RX	MEDLINE=98428793; PubMed=9758550; DOI=10.1016/S0378-1119(
RA	Scott D.A., Grelund J.H. Jr., Marietta J.R., Drury S.,
RA	Sadowski R.E., Vinas A.B., DeAngelis M.M., Cami R., Rames
RA	Kraft M.L., Elbedour K., Skowora A.B., Friedman R.A.,
RA	Stikunaiti Srisalapaty C.R., Verhoeven C., Van Camp G.,
RA	Deininger P.L., Baltzer M.A., Morton C.C., Keats B.J., Smi
RA	Sheffield V.C.;

7	7	64	107	114	151	174
---	---	----	-----	-----	-----	-----

RA Rogers J., Dunham I.;
RT "DNA sequence and analysis of human chromosome 9."; Nature 429:369-374(2004).
RN (3)
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC TISSUE=Eye, Lung, and Spleen;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Bueltow K.H., Schaefer C.F., Bhat N.K., Hopkins S.F., Jordan H., Moore T., Max S.I., Wang J., Heide F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Umedin T.B., Toshiyuki S., Carninci P., Frange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullan S.J., Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butcherfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E., Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences."; Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- TISSUE SPECIFICITY: Expressed in fetal cochlea. Also expressed in infant brain, fetal heart, lung, pancreatic islet, melanocyte, pineal gland, placenta, corneal stroma, and parathyroid tumor.
CC -1- SIMILARITY: Contains 1 A20-type zinc finger.
CC -1- SIMILARITY: Contains 1 AN1-type zinc finger.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.
CC -----
DR EMBL, AF062072; AAC61801.1; -; Genomic_DNA.
DR EMBL, AF062346; AAC42601.1; -; mRNA.
DR EMBL, AF062347; AAC42602.1; -; mRNA.
DR EMBL, AL135924; CAD13440.1; -; Genomic_DNA.
DR EMBL, BC011018; AAH11018.1; -; mRNA.
DR EMBL, BC027707; AAH27707.1; -; mRNA.
DR EMBL, BC073131; AAH73131.1; -; mRNA.
DR SRR, O76080; 140-196.
DR Ensemble, ENSG00000107372; Homo sapiens.
DR HGNC, HGNC:13008; ZA20D2.
DR MIM, 604761; -;
DR InterPro, IPR002653; Znf_A20.
DR InterPro, IPR000058; Znf_AN1.
DR Pfam, PF01754; zF-A20; 1.
DR Pfam, PF01428; zF-AN1; 1.
DR SMART, SM00259; Znf_A20; 1.
DR SMART, SM00154; Znf_AN1; 1.
DR PROSITE, PSS1036; zF_A20; 1.
DR PROSITE, PSS1039; zF_AN1; 1.
DR Metal-binding; Zinc; Zinc-finger.
FT ZN FING 8 42 A20-type.
FT ZN FING 151 194 AN1-type.
SQ SEQUENCE 213 AA; 23132 MW; E477504B1BA77753 CRC64;
Query Match 37.0%; Score 364; DB 1; Length 213;
Best Local Similarity 38.2%; Pred. No. 3.6e-22;
Matches 81; Conservative 31; Mismatches 66; Indels 34; Gaps 7;
QY 6 VSOETTSQAPBGPVWKNLCGPFSGQATMGCSKCYRRTVMQAKTALAEQATQAAQATS 65
DB 1 MNOE-TNQP-CPMLCTGCGFYGNPRITGKCSVCYKELQIQONSGRMPTASGNS 58
QY 66 AT--AAAVO-----PPA--PHETKTLCTCEVTRTIVPHQSS 98

DB 59 PRSDASVQPADTSLNCGAAGSTSEKSRNPVVALPYTQOTKMSI; 118
QY 99 YQODLVTV--PAAAPQAVKSIAPSRPEP--NRGSGCRKRVLTGFI; 153
DB 119 VSEPVVTPSPSVSGSTQSGSEKAPELPKPKNRPFMRKRYGLGFI; 178
QY 154 RYSDKRTCTYDYKAGQEAIAKAPLVVAEKV 185
DB 179 RYSDKNCPCYDYKAEAAAKIRKNPVAEKI 210
RESULT 13
Q6PBK7 BRARE
ID Q6PBK7 BRARE PRELIMINARY; PRT; 213 AA.
AC Q6PBK7;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Zinc finger, A20 domain containing 2.
GN Name=za20d2;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele.
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cyp.
OC Cyprinidae; Danio.
OC NCBI_TaxID=7955;
RN (1)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=M1d-type; TISSUE=Eye;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.24260;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Bueltow K.H., Schaefer C.F., Bhat N.K., Hopkins S.F., Jordan H., Moore T., Max S.I., Wang J., Heide F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences."; Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RT and mouse cDNA sequences."
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN (2)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=M1d-type; TISSUE=Eye;
RC Director MGC Project;
RL Submitted (OCT-2003) to the EMBL/Genbank/DBJ databases.
DR EMBL, BC053673; AAH59673.1; -; mRNA.
DR SRR, Q6PBK7; 140-196.
DR ZFIN, ZDB-GENE-040426-1979; za20d2.
DR GO, GO:0003677; F:DNA binding; IEA.
DR GO, GO:0008270; F:zinc ion binding; IEA.
DR InterPro, IPR002653; Znf_A20.
DR InterPro, IPR000058; Znf_AN1.
DR Pfam, PF01754; zF-A20; 1.
DR Pfam, PF01428; zF-AN1; 1.
DR SMART, SM00259; Znf_A20; 1.
DR SMART, SM00154; Znf_AN1; 1.
SQ SEQUENCE 213 AA; 23280 MW; DC5626C21BA56BA3 CRC64;
Query Match 37.0%; Score 364; DB 2; Length 213;
Best Local Similarity 38.1%; Pred. No. 3.6e-22;
Matches 82; Conservative 33; Mismatches 60; Indels 7;
QY 6 VSOETTSQAPBGPVWKNLCGPFSGQATMGCSKCYRRTVM-----Q 59

```

Db 1 MAOE-TNOSP-V-PILCTTGCGFYGNPRTNGMCSVCYKEHLNRQSSDRSPMSPLAGSPSA 58
QY 60 AAQATSATAAQAQP-----PAPVET-----KLTCEVERTMIIVHOSSSYQQ 101
Db 59 BASAQRLEASINKAETLPAPSETMRBSIPSSLPVQKMT---EMISIREKKALSPXA 115
QY 102 DLVTFAAALAPQAVKSSIAA-----PSPRPNRGSGCRKRVGLTGFRCRGNTYC 150
Db 116 ETVREVITQPTSSYSPPIPAQAANDAKSPDSSKPKKNCFTCRKRVGLTGFRCRGNTYC 175
QY 151 ALHRYSDKHCTTYDYKAAQGAIAKANPLVAAKV 185
Db 176 GHRYSDDKHNCTYDYKABAAAKIRKENPVVADKI 210

```

```

RESULT 14
05ZJR4 CHICK PRELIMINARY; PRT; 212 AA.
AC 05ZJR4;
DT 25-OCT-2004 (TREMBlrel. 28, Created)
DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
DE 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
DE Hypothetical protein.
GN ORFNames=RCJMB04.16e14;
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
RN NCBI_TaxID=9031;
RX [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CB; TISSUE=Bursa;
RA Caldwell R.B., Kierzek A.M., Arakawa H., Bezubov Y., Zaim J.,
RA Fiedler J., Kutter S., Blagoderaki A., Kostovska D., Kotet M.,
RA Plachy J., Carninci P., Hayashizaki Y., Buenstede J.M.,
RT "Full-length cDNAs from chicken bursa lymphocytes to facilitate
RT gene function analysis."
RL Genome Biol. 6:R6-R6(2005).
DR EMBL, AJ720370; CAC32029.1; -, mRNA.
DR SMR, Q5ZJR4; 139-195.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR InterPro; IPR002653; Znf_A20.
DR InterPro; IPR000058; Znf_AN1.
DR Pfam; PF01428; ZF-A20; 1.
DR Pfam; PF01428; ZF-A20; 1.
DR SMART; SM00259; Znf_A20; 1.
DR SMART; SM00154; Znf_AN1; 1.
DR Hypothetical protein.
KW Hypothetical protein.
SQ SEQUENCE 212 AA; 23141 MW; D974C13CF9B935A9 CRC64;

```

```

Query Match 36.9%; Score 363.5; DB 2; Length 212;
Best Local Similarity 38.0%; Pred. No. 3,9e-22;
Matches 84; Conservative 27; Mismatches 57; Indels 53; Gaps 8;
QY 6 VSQETTSQAPBGPVWCKNLGFPFGSOATWGLCSKCYR----- 43
Db 1 MTQE-TNCTP-GPMICSTGCGFYGNPRTNGMCSVCYKEHLNRQSSDRSPMSPLAGSPSA 58
QY 44 -----TWQAKMTALAEQAQTAQAQTSATAAQAQP-----HETKLTCV 86
Db 59 PTSBSTSVQADTSL-NNCDGAVGSTSEKSRSPVVALPVQQTMEISIREKVTPT 117
QY 87 ERTMIVPHOSSSYQODLVTPAAAAPQAVKSSIAAPSRP-----NRGSGCRKRVGLTGFRCR 144
Db 118 ETEPVVTP-----TPSVQSPSTSRNEKAPELPKPKKRCFMCKRVGLTGFDCR 168
QY 145 CGNLFCALHRYSDKHCTTYDYKAAQGAIAKANPLVAAKV 185
Db 169 CGNLFCGLHRYSDKHCTTYDYKABAAAKIRKENPVVADKI 209

```

RESULT 15

```

Q9LX15 ARATH
ID Q9LX15 ARATH PRELIMINARY; PRT; 170 AA.
AC Q9LX15;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Zinc finger-1like protein.
GN Name=ZF22.200;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Trach
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicot
OC eucotsids II; Brassicales; Brassicaceae; Arabidopsids.
OC NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Purnelle B., Masny D., Goffeau A., Boutry M., Mewes H.W.,
RA Lemcke K., Mayer K.F.X., Quetier F., Salanoubat M.;
RL Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA EU Arabidopsis sequencing project;
RL Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL353912; CAB89241.1; -, Genomic DNA.
DR PIR; T49033; T49033.
DR SMR; Q9LX15; 107-151.
DR InterPro; IPR002653; Znf_A20.
DR InterPro; IPR000058; Znf_AN1.
DR Pfam; PF01428; ZF-A20; 1.
DR SMART; SM00154; Znf_AN1; 1.
SQ SEQUENCE 170 AA; 18598 MW; 81B28298CF83D620 CRC64;

```

```

Query Match 36.8%; Score 362.5; DB 2; Length 170;
Best Local Similarity 37.8%; Pred. No. 3.8e-22;
Matches 73; Conservative 32; Mismatches 53; Indels 4;
QY 6 VSQETTSQAPBGPVWCKNLGFPFGSOATWGLCSKCYRTVWQAKMTALA 65
Db 1 MAEHRQCTPESNRNLGVNNGCFLGSSATMNLCSNCYGDLCK----- 54
QY 66 ATAAVQPPAVVHETKLTCVEVERTMIIVHOSSSYQODLVTPAAAAPQA- 119
Db 55 ESSISVSPPS-----SSSSEISSPIIPPLKNPSVK 96
QY 120 APS-----RPEPNRGSGCRKRVGLTGFRCRGNTYCALHRYSDKHCTTY 174
Db 97 LPTTQNOQQRPNKCTTCKRVGLTGFRCRCGTMCVGHRYPEIHGCSY 156
QY 175 KANPLVAAKVVK 187
Db 157 KANPLVKAALKOK 169

```

Search completed: December 8, 2005, 16:32:01
Job time : 162 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 2005, 15:03:00 ; Search time 44 Seconds
(without alignments)
353.251 Million cell updates/sec

Title: US-10-716-089-20

Perfect score: 984
Sequence: 1 MATEVSOEITTSQAEPEGVPM.....GGEAIKANPLVAEKVKVF 188

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 200000000Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/pdata/1/1aa/5_COMB.pep:*
2: /cgn2_6/pdata/1/1aa/6_COMB.pep:*
3: /cgn2_6/pdata/1/1aa/H_COMB.pep:*
4: /cgn2_6/pdata/1/1aa/ECTUS_COMB.pep:*
5: /cgn2_6/pdata/1/1aa/RE_COMB.pep:*
6: /cgn2_6/pdata/1/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	984	100.0	188	2	US-09-828-303-20 Sequence 20, Appl
2	366	37.2	213	2	US-09-244-805-11 Sequence 11, Appl
3	365	37.1	395	2	US-09-949-016-9564 Sequence 9564, Ap
4	364	37.0	213	2	US-09-949-016-6681 Sequence 6681, Ap
5	343.5	34.9	208	1	US-08-861-269-3 Sequence 3, Appl
6	343.5	34.9	208	1	US-09-134-596-3 Sequence 3, Appl
7	343.5	34.9	208	2	US-09-293-273-3 Sequence 3, Appl
8	313	31.8	189	1	US-08-861-269-7 Sequence 7, Appl
9	313	31.8	189	1	US-09-134-596-7 Sequence 7, Appl
10	313	31.8	189	2	US-09-293-273-7 Sequence 7, Appl
11	177	18.0	46	1	US-08-861-269-8 Sequence 8, Appl
12	177	18.0	46	1	US-09-134-596-8 Sequence 8, Appl
13	177	18.0	46	2	US-09-293-273-8 Sequence 8, Appl
14	101	10.3	497	2	US-09-134-000C-5990 Sequence 5990, Ap
15	99.5	10.1	491	1	US-08-942-819-2 Sequence 2, Appl
16	98.5	10.1	491	2	US-09-522-955A-2 Sequence 2, Appl
17	98	10.0	87	2	US-09-513-999C-5919 Sequence 5919, Ap
18	98	10.0	258	2	US-10-012-231A-153 Sequence 153, App
19	98	10.0	258	2	US-10-015-389A-153 Sequence 153, App
20	98	10.0	258	2	US-10-006-768A-153 Sequence 153, App
21	98	10.0	258	2	US-10-015-671A-153 Sequence 153, App
22	98	10.0	258	2	US-10-015-393A-153 Sequence 153, App
23	98	10.0	258	2	US-10-011-833A-153 Sequence 153, App
24	98	10.0	258	2	US-10-006-041A-153 Sequence 153, App
25	98	10.0	258	2	US-10-012-064A-153 Sequence 153, App
26	92.5	9.4	480	2	US-09-108-020-6 Sequence 6, Appl
27	92.5	9.4	480	2	US-09-685-296-6 Sequence 6, Appl

28	92	9.3	414	2	US-09-270-767-57349	Sec	' A
29	92	9.3	506	2	US-09-270-767-42086	Sec	' A
30	89	9.0	174	2	US-09-248-796A-25211	Sec	' A
31	88	8.9	595	1	US-08-232-087A-2	Sec	' A
32	88	8.9	595	2	US-09-006-353A-9	Sec	' A
33	88	8.9	595	2	US-09-573-986-9	Sec	' A
34	88	8.9	595	2	US-09-949-016-6048	Sec	' A
35	88	8.9	642	2	US-09-949-016-8043	Sec	' A
36	87.5	8.9	176	2	US-10-104-047-2419	Sec	' A
37	87	8.8	595	1	US-08-570-923-2	Sec	' A
38	87	8.8	595	1	US-08-225-989-2	Sec	' A
39	87	8.8	595	1	US-08-580-014-2	Sec	' A
40	87	8.8	595	2	US-09-079-785-2	Sec	' A
41	87	8.8	595	2	US-09-921-667-6	Sec	' A
42	86.5	8.8	595	2	US-09-628-126-2	Sec	' A
43	86.5	8.8	361	2	US-09-976-594-351	Sec	' A
44	85.5	8.7	456	2	US-09-689-486-56	Sec	' A
45	84.5	8.6	792	2	US-09-949-016-10798	Sec	' A

ALIGNMENTS

```
RESULT 1
US-09-828-303-20
; Sequence 20, Application US/09828303
; Patent No. 6677504
; GENERAL INFORMATION:
; APPLICANT: COSTA E SILVA, OSWALDO DA
; APPLICANT: BOHNETT, HANS J.
; APPLICANT: VAN THIELEN, NOCHA
; APPLICANT: CHEN, ROUYING
; TITLE OF INVENTION: TRANSCRIPTION FACTOR STRESS-RELATED PRO.
; FILE REFERENCE: 16313-0030
; CURRENT APPLICATION NUMBER: US/09/828,303
; CURRENT FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: 60/196,001
; PRIOR FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Physcomitrella patens
US-09-828-303-20

Query Match          100.0%; Score 984; DB 2; Length 1
Best Local Similarity 100.0%; Pred. No. 1e-97;
Matches 188; Conservative 0; Mismatches 0; Indels 0;

QY      1 MATEVSOEITTSQAEPEGVPMCKNLGFFGSOATMGICSKCYRETYMOM
DB      1 MATEVSOEITTSQAEPEGVPMCKNLGFFGSOATMGICSKCYRETYMOM
QY      61 AQAISATPAAVOPAPVHETLTCEVERTWIVPHOSSSYOODLVTPAAV
DB      61 AQAISATPAAVOPAPVHETLTCEVERTWIVPHOSSSYOODLVTPAAV
QY      121 PSRPPNRCGSCRKRVGLTGFKRCGNLYCALHYSYDQHTTYKKAQ
DB      121 PSRPPNRCGSCRKRVGLTGFKRCGNLYCALHYSYDQHTTYKKAQ
QY      181 VAEKVKVF 188
DB      181 VAEKVKVF 188

RESULT 2
US-09-244-805-11
; Sequence 11, Application US/09244805
; Patent No. 6699660
; GENERAL INFORMATION:
```

APPLICANT: Worley, Paul F.
APPLICANT: Lanahan, Anthony
APPLICANT: Goetz, Bernard
APPLICANT: Helmisch, Holger
APPLICANT: Kuner, Rohnd
APPLICANT: Scheek, Sigrid
APPLICANT: Nikolich, Karoly
APPLICANT: Zhukovskii, Eugene
TITLE OF INVENTION: IMMEDIATE EARLY GENES AND METHODS OF USE
FILE REFERENCE: 10496/004001
CURRENT FILING DATE: 1999-02-05
PRIOR FILING DATE: 1998-02-12
PRIOR APPLICATION NUMBER: 60/074,518
PRIOR FILING DATE: 1998-02-12
PRIOR APPLICATION NUMBER: 60/074,135
NUMBER OF SEQ ID NOS: 62
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11
LENGTH: 213
TYPE: PRT
ORGANISM: Eukaryote
US-09-244-805-11

Query Match 37.2%; Score 366; DB 2; Length 213;
Best Local Similarity 38.2%; Pred. No. 2.8e-31;
Matches 81; Conservative 32; Mismatches 65; Indels 34; Gaps 7;

6 VSOETTSGAPBGPVWCKNLGFFGSOATMGLCSKCYRETVMQAKMTALAEOATQAQATKS 65
1 MAQE-TNQTG-PGMLCSTGCGFYGNPRTMGCSVCYKHLRQNGSGMSPGTAQSGNS 58
66 AT-AAAVQ-----PPA--PVHETKLTCEVETMIVHQS 98
59 PTSDASVQADATLNCEGAAGSTSEKSRNVPAALPTQMTMSISREDKITSPKTE 118
99 YQODLVLT---PAAAPQAVKSIAPSRPEP--NRGSCRKRVGLTGFPCRCGNLYCALH 153
119 VSEPVVTPSPSVQSPSSQSEKAPBELPKPKKNCFCMKRKVGLTGDFCRGNLFCGLH 178
154 RYSDKHTCTYDYKAAQGEALAKANPLVVAEKV 185
179 RYSDKNCPPYDYKAAAKIRKENPVVAAEKI 210

RESULT 3

US-09-949-016-9564
Sequence 9564, Application US/0949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: CLO01307
CURRENT FILING DATE: 2000-04-14
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 9564
LENGTH: 395
TYPE: PRT
ORGANISM: Human
US-09-949-016-9564

Query Match 37.1%; Score 365; DB 2; Length 395;
Best Local Similarity 37.9%; Pred. No. 8.1e-31;

Matches 81; Conservative 32; Mismatches 67; Indels 7;
4 ERVSOETTSGAPBGPVWCKNLGFFGSOATMGLCSKCYRETVMQAKMTA 63
181 KMAQE-TNQTG-PGMLCSTGCGFYGNPRTMGCSVCYKHLRQNGSG 238
64 TSAT-AAAVQ-----PPA--PVHETKLTCE 96
239 NSPTSDASVQADATLNCEGAAGSTSEKSRNVPAALPTQMTMS 298
97 SSVQODLVLT---PAAAPQAVKSIAPSRPEP--NRGSCRKRVGLTG 151
299 TEVSEPVVTPSPSVQSPSSQSEKAPBELPKPKKNCFCMKRKVGLTG 358
152 LHRYSKHTCTYDYKAAQGEALAKANPLVVAEKV 185
359 LHRYSKNCPPYDYKAAAKIRKENPVVAAEKI 392

RESULT 4

US-09-949-016-6681
Sequence 6681, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: CLO01307
CURRENT FILING DATE: 2000-04-14
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6681
LENGTH: 213
TYPE: PRT
ORGANISM: Human
US-09-949-016-6681

HEREC

Query Match 37.0%; Score 364; DB 2; Length 213
Best Local Similarity 38.2%; Pred. No. 4.5e-31;
Matches 81; Conservative 31; Mismatches 66; Indels 7;

6 VSOETTSGAPBGPVWCKNLGFFGSOATMGLCSKCYRETVMQAKMTALA 65
1 MAQE-TNQTG-PGMLCSTGCGFYGNPRTMGCSVCYKHLRQNGSGM 58
66 AT-AAAVQ-----PPA--PVHETKLTCEV 98
59 PTSDASVQADATLNCEGAAGSTSEKSRNVPAALPTQMTMSIS 118
99 YQODLVLT---PAAAPQAVKSIAPSRPEP--NRGSCRKRVGLTGK 153
119 VSEPVVTPSPSVQSPSSQSEKAPBELPKPKKNCFCMKRKVGLTG 178
154 RYSDKHTCTYDYKAAQGEALAKANPLVVAEKV 185
179 RYSDKNCPPYDYKAAAKIRKENPVVAAEKI 210

RESULT 5

US-08-861-269-3
Sequence 3, Application US/08861269
Patent No. 5817494
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Corley, Neil C.
APPLICANT: Lal, Preeti

```

1 APPLICANT: Shah, Purvi
2 TITLE OF INVENTION: UBIQUITIN CONJUGATION PROTEINS
3 NUMBER OF SEQUENCES: 8
4 CORRESPONDENCE ADDRESS:
5 ADDRESSEE: Incyte Pharmaceuticals, Inc.
6 STREET: 3174 Porter Drive
7 CITY: Palo Alto
8 STATE: CA
9 COUNTRY: USA
10 ZIP: 94304
11 COMPUTER READABLE FORM:
12 MEDIUM TYPE: Diskette
13 COMPUTER: IBM Compatible
14 OPERATING SYSTEM: DOS
15 SOFTWARE: PseSeq for Windows Version 2.0
16 CURRENT APPLICATION DATA:
17 APPLICATION NUMBER: US/08/861,269
18 FILING DATE: Filed Herewith
19 PRIOR APPLICATION DATA:
20 APPLICATION NUMBER:
21 FILING DATE:
22 ATTORNEY/AGENT INFORMATION:
23 NAME: Billings, Lucy J.
24 REGISTRATION NUMBER: 36,749
25 REFERENCE/DOCKET NUMBER: PP-0302 US
26 TELECOMMUNICATION INFORMATION:
27 TELEPHONE: 415-855-0555
28 TELEFAX: 415-845-4166
29 INFORMATION FOR SEQ ID NO: 3:
30 SEQUENCE CHARACTERISTICS:
31 LENGTH: 208 amino acids
32 TYPE: amino acid
33 STRANDEDNESS: single
34 TOPOLOGY: linear
35 IMMEDIATE SOURCE:
36 LIBRARY: BRAINOT09
37 CLONE: 2151473
38
39 US-08-861-269-3
40
41 Query Match 34.9%; Score 343.5; DB 1; Length 208;
42 Best Local Similarity 36.4%; Pred No. 7e-29;
43 Matches 78; Conservative 31; Mismatches 66; Indels 39; Gaps 7.
44
45 QY 6 VSQETTSQAPBGPVWCXNLCGPFSGSQATWGLCSKCYRE-----TWQAK 49
46 Db 1 MAQENIHNS--GVPMLCSTGCCFGYGPRITNGMGSVCYKHHLGRQNSNGRISPPATSLSL 58
47
48 QY 50 MTALAEQAT-----QAQOTSATAAIVOPPAHVHETKLTCVEERTMIVIPHOSSYYQQDLY 104
49 Db 59 SESLPVQCCTDSVPBAQSALDSTSSMO--PSFVSNQSLISE----SVASSQLDSTSVDKA 113
50
51 QY 105 TPAAAPAAP-VKSSLAAPS-----RPDPNCGSCRRKRGVLTKPKRCGNLYCALH 153
52 Db 114 VPETEDVASVSDTAQQPSEESQSKLEKPKOKKNCFMCCKRKVGILGFECRCGNVYCGVH 173
53
54 QY 154 RYSDKHTCTDYDKAGOEAIKANPLVVAAEKVV 187
55 Db 174 RYSDVHNCSYNKKADAERIKRENVTVGEEKIQK 207
56
57 RESULT 6
58 US-09-134-596-3
59 Sequence 3, Application US/09134596
60 Patent No. 5923318
61 GENERAL INFORMATION:
62 APPLICANT: Bandman, Olga
63 APPLICANT: Corley, Neil C.
64 APPLICANT: Lal, Preeti
65 APPLICANT: Shah, Purvi
66 TITLE OF INVENTION: UBIQUITIN CONJUGATION PROTEINS
67 NUMBER OF SEQUENCES: 8
68 CORRESPONDENCE ADDRESS:
69 ADDRESSEE: Incyte Pharmaceuticals, Inc.
```

1 STREET: 3174 Porter Drive
2
3 CITY: Palo Alto
4 STATE: CA
5 COUNTRY: USA
6 ZIP: 94304
7
8 COMPUTER READABLE FORM:
9 MEDIUM TYPE: Diskette
10 COMPUTER: IBM Compatible
11 OPERATING SYSTEM: DOS
12 SOFTWARE: FASTSEQ for Windows Version 2.0
13
14 CURRENT APPLICATION DATA:
15 APPLICATION NUMBER: US/09/134,596
16
17 FILING DATE:
18
19 PRIORITY APPLICATION DATA:
20 APPLICATION NUMBER: 08/861,269
21
22 FILING DATE:
23
24 ATTORNEY/AGENT INFORMATION:
25
26 NAME: Billings, Lucy J.
27 REGISTRATION NUMBER: 36,749
28 REFERENCE/DOCKET NUMBER: PF-0302 US
29
30 TELECOMMUNICATION INFORMATION:
31
32 TELEPHONE: 415-855-0555
33
34 TELEFAX: 415-845-4166
35
36 INFORMATION FOR SEQ ID NO: 3:
37
38 SEQUENCE CHARACTERISTICS:
39
40 LENGTH: 208 amino acids
41 TYPE: amino acid
42 STRANDEDNESS: single
43 TOPOLOGY: linear
44
45 IMMEDIATE SOURCE:
46
47 LIBRARY: BRAINOT09
48
49 CLONE: 2151473
50
51 US-09-134-596-3

Query Match	34.9%; Score 343.5; DB 1; Length :	
Beat Local Similarity	36.4%; Pred. No. 7e-29;	
Matches	78; Conservative 31; Mismatches 66; Indels	7
QY	6 VSOETTSQAEQGPVWCNKLGSFPGSQATMGLGCKCYRE-----	49
	:: : :: ::	
Db	1 MAQGTNHS--QVPMLCSTGCGFYGNPRTNGMCSICYKEHLORONSNGF	58
QY	50 MTLALAQAT-----QAQAQTSATAAAVQAPAPVHETKLTCVEKRTMIVF	104
	:: : :: ::	
Db	59 SESLPVQCTDGSVPEAQSALDSTSSMQ-PSPVSNOSLSB-----SVAS	113
QY	105 TPAALAAPOA-VKSSIAAPS-----RPEPRQSCSRKRVGLTGPK	153
Db	114 VPETEDVQASVSDTAQCPSEBSKSLKPKQKQRKRCMKRKGVLGTPE	173
QY	154 RYSDKHTCTDYKKAAGEATAKANPLVAEKLVVK 187	
Db	174 RYSDVNCSTNYKADAAEKIRKENPVVVGKEIKQK 207	

RESULT 7
US-09-293-273-3

: Sequence 3, Application US/09293273

: Patent No. 6057112

: GENERAL INFORMATION:

: APPLICANT: Bandman, Olga

: APPLICANT: Corley, Neil C.

: APPLICANT: Lal, Preeti

: APPLICANT: Shah, Purvi

: TITLE OF INVENTION: UBIQUITIN CONJUGATION PROTEINS

: NUMBER OF SEQUENCES: 8

: CORRESPONDENCE ADDRESS:

: ADDRESSEE: Incyte Pharmaceuticals, Inc.

: STREET: 3174 Porter Drive

: CITY: Palo Alto

: STATE: CA

: COUNTRY: USA

: ZIP: 94304


```
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/293,273
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/134,596
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0302 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 208 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BRAINOT09
CLONE: 2151473
US-09-293-273-3
```

```
Query Match          34.9%; Score 343.5; DB 2; Length 208;
Best Local Similarity 36.4%; Pred. No. 7e-29;
Matches 78; Conservative 31; Mismatches 66; Indels 39; Gaps 7;

QY 6 VSQETSAPEBPVCKNLGFFGSOATMGLCSKYRRE-----TWQAK 49
   :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
DB 1 MAQETNHS--QVPMLCSTGCGFYGNPRTGMCVCYKEHLQRONSSNGRISPPATSVSSL 58
   :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
QY 50 MPALEQAT-----QAAQATSAATAAIVOPAPVHETKLTCEERTMIVPHOSSVQQDL 104
   :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
DB 59 SSSLVQCTDGSVPEKQSLDSTSSMQ-PSPVSNOSLSL-----SVASSQLDSTSDKA 113
   :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
QY 105 TPAAAPQA-VKSIAPG-----RPPNRCGSGCRKRVGLTGFKRCGNLYCALH 153
   :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
DB 114 VPEDEDVQASVSDTAQQPEEQSKSLKPKQKQKRCFMKRVGLTGFECRCGNYCGVH 173
   :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
QY 154 RYSDKHTCTYDKAAGEIAKAPLYVAEKYVK 187
   :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
DB 174 RYSDVHNCSTYNYKADAEXIKRENPNVVGKIQK 207
   :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
```

```
RESULT 8
US-08-861-269-7
Sequence 7, Application US/08861269
Patent No. 5817494
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Corley, Neil C.
APPLICANT: Lal, Preeti
APPLICANT: Shah, Purvi
TITLE OF INVENTION: UBIQUITIN CONJUGATION PROTEINS
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSER: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
```

```
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/861,269
FILING DATE: Filed Herewith
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0302 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 189 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 1279278
US-08-861-269-7
```

```
Query Match          31.8%; Score 313; DB 1; Length 189
Best Local Similarity 35.4%; Pred. No. 1.2e-25;
Matches 68; Conservative 31; Mismatches 69; Indels 6;

QY 8 QETTSQAPEBPVCKNLGFFGSOATMGLCSKYRETYMAKMTALAEQ 67
   :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
DB 5 QQAQTAAPS-----CRAGGFFGASATBEGYCSQCFKNTLKRQODTVRLTS 60
   :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
QY 68 AAAYQ-PPAPV-----HETKLTCEERTMIVPHOSSVQQDL 113
   :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
DB 61 SSALKSEPSYDMCKMAVSDETAQKDCG-----DIINVCDQINDS 115
   :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
QY 114 VKSIIAPSRPEPNRCGSGCRKRVGLTGFKRCGNLYCALHRYSDKHTCT 173
   :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
DB 116 ITVDVPVPVK-KANRCHMKKRVGLTGFSCRCGGLYCGDHHYDQAHCNC 174
   :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
QY 174 AKANPLVVAERY 185
   :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
DB 175 RKNPNVVVSDKV 186
   :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
```

```
RESULT 9
US-09-134-596-7
Sequence 7, Application US/09134596
Patent No. 5922318
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Corley, Neil C.
APPLICANT: Lal, Preeti
APPLICANT: Shah, Purvi
TITLE OF INVENTION: UBIQUITIN CONJUGATION PROTEINS
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSER: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/134,596
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/861,269
```



```

; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0302 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 189 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 1279278
; US-09-134-596-7

Query Match      31.8%; Score 313; DB 1; Length 189;
Best Local Similarity 35.4%; Pred. No. 1.2e-25;
Matches 68; Conservative 31; Mismatches 69; Indels 24; Gaps 6;

QY 8 QETSOAPBGPVMCKNLGCFPGSOATMGLCSKCYRETVMQAKMTALAEQATQAQATSAT 67
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 5 QQAQTAAPS-----CRAGCGFFGASATBEGYCSQCFTKTLRKQDITVRLTSPVSPSSMAAT 60
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

QY 68 AAAPV-PPAPV-----HETKLTCEVERTMIVPHOSSSYQODLVTPA-AAAPQA 113
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 61 SALSSEPSVDMCKKAIVSDETAAMDCE-----DIINVDDQINDBSVTVAESTAPTT 115
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

QY 114 VKSSTIAPSRPEPNRCGSKRRVGLTGFCRCGNLYCALHRSYDKTCTYDYKAAGQEA 173
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 116 ITVDVPEVPK-KANRCHMKCKRRVGLTGFCRCGGLYCGDHRDYDQAHNCC 174
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

QY 174 AKANPLVVAEKY 185
DB 175 KKNNPVVSDKV 186

RESULT 10
US-09-293-273-7
; Sequence 7, Application US/09293273
; Patent No. 6057112
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Corley, Neil C.
; APPLICANT: Lal, Preeti
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: UB1QUTIN CONJUGATION PROTEINS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/293,273
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/134,596
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0302 US
```

```

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 189 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 1279278
; US-09-293-273-7

Query Match      31.8%; Score 313; DB 2; Length 189;
Best Local Similarity 35.4%; Pred. No. 1.2e-25;
Matches 68; Conservative 31; Mismatches 69; Indels 6;

QY 8 QETSOAPBGPVMCKNLGCFPGSOATMGLCSKCYRETVMQAKMTALAE 67
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 5 QQAQTAAPS-----CRAGCGFFGASATBEGYCSQCFTKTLRKQDITVRLT 60
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

QY 68 AAAPV-PPAPV-----HETKLTCEVERTMIVPHOSSSYQODLV 113
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 61 SALSSEPSVDMCKKAIVSDETAAMDCE-----DIINVDDQINDBS 115
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

QY 114 VKSSTIAPSRPEPNRCGSKRRVGLTGFCRCGNLYCALHRSYDKTCT 173
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 116 ITVDVPEVPK-KANRCHMKCKRRVGLTGFCRCGGLYCGDHRDYDQAHNCC 174
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

QY 174 AKANPLVVAEKY 185
DB 175 KKNNPVVSDKV 186

RESULT 11
US-08-861-269-8
; Sequence 8, Application US/08861269
; Patent No. 5817494
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Corley, Neil C.
; APPLICANT: Lal, Preeti
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: UB1QUTIN CONJUGATION PROTEINS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/861,269
; FILING DATE: Filed Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0302 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
```

```
/ LENGTH: 46 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ IMMEDIATE SOURCE:
/ LIBRARY: GenBank
/ CLONE: 998680
US-08-861-269-8

Query Match 18.0%; Score 177; DB 1; Length 46;
Best Local Similarity 69.8%; Pred. No. 7e-12;
Matches 30; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 143 CRCGNLYCALHRYSDKHCTCTYYKAGGEAIAKAPLVVAEKV 185
DB 1 CRCGNLYCALHRYSDKHCTCTYYKAGGEAIAKAPLVVAEKI 43

RESULT 12
US-09-134-596-8
/ Sequence 8, Application US/09134596
/ Patent No. 592318
/ GENERAL INFORMATION:
/ APPLICANT: Bandman, Olga
/ APPLICANT: Corley, Neil C.
/ APPLICANT: Lal, Preeti
/ APPLICANT: Shah, Purvi
/ TITLE OF INVENTION: UBIQUITIN CONJUGATION PROTEINS
/ NUMBER OF SEQUENCES: 8
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Incyte Pharmaceuticals, Inc.
/ STREET: 3174 Porter Drive
/ CITY: Palo Alto
/ STATE: CA
/ COUNTRY: USA
/ ZIP: 94304
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: DOS
/ SOFTWARE: FASTSEQ for Windows Version 2.0
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/134,596
/ FILING DATE:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/861,269
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Billings, Lucy J.
/ REGISTRATION NUMBER: 36,749
/ REFERENCE/DOCKET NUMBER: PF-0302 US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 415-855-0555
/ TELEFAX: 415-845-4166
/ INFORMATION FOR SEQ ID NO: 8:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 46 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ IMMEDIATE SOURCE:
/ LIBRARY: GenBank
/ CLONE: 998680
US-09-134-596-8

Query Match 18.0%; Score 177; DB 1; Length 46;
Best Local Similarity 69.8%; Pred. No. 7e-12;
Matches 30; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 143 CRCGNLYCALHRYSDKHCTCTYYKAGGEAIAKAPLVVAEKV 185
DB 1 CRCGNLYCALHRYSDKHCTCTYYKAGGEAIAKAPLVVAEKI 43
```

```
RESULT 13
US-09-293-273-8
/ Sequence 8, Application US/09293273
/ Patent No. 6057112
/ GENERAL INFORMATION:
/ APPLICANT: Bandman, Olga
/ APPLICANT: Corley, Neil C.
/ APPLICANT: Lal, Preeti
/ APPLICANT: Shah, Purvi
/ TITLE OF INVENTION: UBIQUITIN CONJUGATION PROTEINS
/ NUMBER OF SEQUENCES: 8
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Incyte Pharmaceuticals, Inc.
/ STREET: 3174 Porter Drive
/ CITY: Palo Alto
/ STATE: CA
/ COUNTRY: USA
/ ZIP: 94304
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: DOS
/ SOFTWARE: FASTSEQ for Windows Version 2.0
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/293,273
/ FILING DATE:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 09/134,596
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Billings, Lucy J.
/ REGISTRATION NUMBER: 36,749
/ REFERENCE/DOCKET NUMBER: PF-0302 US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 415-855-0555
/ TELEFAX: 415-845-4166
/ INFORMATION FOR SEQ ID NO: 8:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 46 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ IMMEDIATE SOURCE:
/ LIBRARY: GenBank
/ CLONE: 998680
US-09-293-273-8

Query Match 18.0%; Score 177; DB 2; Length 46;
Best Local Similarity 69.8%; Pred. No. 7e-12;
Matches 30; Conservative 3; Mismatches 10; Indels 0;

QY 143 CRCGNLYCALHRYSDKHCTCTYYKAGGEAIAKAPLVVAEKV 185
DB 1 CRCGNLYCALHRYSDKHCTCTYYKAGGEAIAKAPLVVAEKI 43

RESULT 14
US-09-134-000C-5990
/ Sequence 5990, Application US/09134000C
/ Patent No. 6617156
/ GENERAL INFORMATION:
/ APPLICANT: Lynn Doucette-Stamm et al
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RE
/ FILE REFERENCE: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS A
/ CURRENT FILING DATE: 1998-08-13
/ PRIOR FILING DATE: 1997-08-15
/ NUMBER OF SEQ ID NOS: 6812
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 5990
```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 2005, 16:26:56 ; Search time 116 Seconds

(without alignments)
677.172 Million cell updates/sec

Title: US-10-716-089-20

Perfect score: 984
Sequence: 1 MATEVSOETTSGAPEGVPM.....GGEALAKANPLVAKVKYKF 188

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA Main:
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBSCOMB.rep:*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBSCOMB.rep:*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBSCOMB.rep:*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBSCOMB.rep:*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBSCOMB.rep:*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBSCOMB.rep:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	984	100.0	188	US-09-828-303-20	Sequence 20, Appl
2	984	100.0	188	US-10-716-089-20	Sequence 20, Appl
3	425.5	43.2	171	US-10-767-701-47138	Sequence 47138, A
4	419.5	42.6	171	US-10-437-963-123596	Sequence 123596, A
5	416.5	42.3	173	US-10-437-963-140292	Sequence 140292, A
6	416.5	42.3	338	US-10-437-963-140292	Sequence 140292, A
7	415	42.2	193	US-10-425-115-261492	Sequence 261492, A
8	415	42.2	211	US-10-425-115-261492	Sequence 261492, A
9	415	42.2	224	US-10-425-115-261494	Sequence 261494, A
10	414.5	42.1	171	US-10-425-115-256094	Sequence 256094, A
11	414.5	41.9	179	US-10-425-114-59729	Sequence 59729, A
12	412.5	41.9	179	US-10-425-114-65423	Sequence 65423, A
13	411.5	41.8	171	US-10-767-701-45701	Sequence 45701, A
14	407.5	41.4	171	US-10-437-963-123595	Sequence 123595, A
15	406.5	41.3	174	US-10-425-115-308732	Sequence 308732, A
16	403	41.0	205	US-10-425-115-256116	Sequence 256116, A
17	402.5	40.9	173	US-10-425-114-44878	Sequence 44878, A
18	402.5	40.9	173	US-10-425-114-57001	Sequence 57001, A
19	401.5	40.8	171	US-10-425-115-185981	Sequence 185981, A
20	401.5	40.8	171	US-10-425-115-185983	Sequence 185983, A
21	401.5	40.8	172	US-10-424-599-178876	Sequence 178876, A
22	401.5	40.8	173	US-10-425-114-57676	Sequence 57676, A
23	395.5	40.2	164	US-10-424-599-231901	Sequence 231901, A
24	392.5	39.9	159	US-10-425-115-270156	Sequence 270156, A
25	391.5	39.8	160	US-10-424-599-231898	Sequence 231898, A
26	391.5	39.8	160	US-10-424-599-231899	Sequence 231899, A
27	390	39.6	247	US-10-425-115-268773	Sequence 268773, A

28	389	39.5	174	US-10-424-599-204286	Seq	3
29	386	39.2	160	US-10-437-963-154553	Seq	3
30	386	39.2	172	US-10-767-701-44663	Seq	3
31	384.5	39.1	161	US-10-437-963-112566	Seq	3
32	383.5	39.0	175	US-10-424-599-204285	Seq	3
33	374.5	38.1	172	US-10-424-599-203681	Seq	3
34	374	38.0	172	US-10-424-599-242209	Seq	3
35	366	37.2	213	US-09-244-805-11	Seq	9
36	366	37.2	213	US-09-245-277-11	Seq	9
37	366	37.2	213	US-10-792-481-11	Seq	9
38	365.5	37.1	172	US-10-424-599-242210	Seq	9
39	365	37.1	229	US-10-264-049-2758	Seq	9
40	365	37.1	310	US-09-925-300-1344	Seq	9
41	364	37.0	213	US-10-507-617-27	Seq	9
42	361	36.7	224	US-10-437-963-116025	Seq	9
43	357	36.3	163	US-10-425-115-369121	Seq	5
44	357	36.3	224	US-10-425-115-71695	Seq	5
45	355.5	36.1	198	US-10-425-115-256050	Seq	0

ALIGNMENTS

```

RESULT 1
US-09-828-303-20
; Sequence 20, Application US/09828303
; Patent No. US20020102695A1
; GENERAL INFORMATION:
; APPLICANT: COSTA E SILVA, OSWALDO DA
; APPLICANT: BOHNETT, HANS J.
; APPLICANT: VAN THIELEN, NOCHIA
; APPLICANT: CHEN, ROUYING
; TITLE OF INVENTION: TRANSCRIPTION FACTOR STRESS-RELATED PRO.
; FILE REFERENCE: 16313-0030
; CURRENT APPLICATION NUMBER: US/09/828,303
; CURRENT FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: 60/196,001
; PRIOR FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Physcomitrella patens
US-09-828-303-20

Query Match      100.0%; Score 984; DB 3; Length 18
Best Local Similarity 100.0%; Pred. NO. 6.7e-85;
Matches 188; Conservative 0; Mismatches 0; Indels 0;

QY      1 MATEVSOETTSGAPEGVPMCKNLGFGSOATMGLCSKCYRETVMQAK
DB      1 MATEVSOETTSGAPEGVPMCKNLGFGSOATMGLCSKCYRETVMQAK
QY      61 AQAISATAAAVQPPAPVHETKLTCEVERTWIVPHOSSYYODLVTPAAV
DB      61 AQAISATAAAVQPPAPVHETKLTCEVERTWIVPHOSSYYODLVTPAAV
QY      121 PSRPPNRCGSKRKVGLTGPKRCGNNLYCALHRYSDDHCTYDYKAK
DB      121 PSRPPNRCGSKRKVGLTGPKRCGNNLYCALHRYSDDHCTYDYKAK
QY      181 VAEKVKVF 188
DB      181 VAEKVKVF 188
QY      181 VAEKVKVF 188
DB      181 VAEKVKVF 188

RESULT 2
US-10-716-089-20
; Sequence 20, Application US/10716089
; Publication No. US20040107463A1
; GENERAL INFORMATION:

```

APPLICANT: COSTA E SILVA, OSWALDO DA
APPLICANT: BOHNETT, HANS J.
APPLICANT: VAN THIELEN, NOCHA
APPLICANT: CHEN, KOUYING
TITLE OF INVENTION: TRANSCRIPTION FACTOR STRESS-RELATED PROTEINS AND
TITLE OF INVENTION: METHODS OF USE IN PLANTS
FILE REFERENCE: 16313-0030
CURRENT APPLICATION NUMBER: US/10/716,089
CURRENT FILING DATE: 2003-11-18
PRIOR APPLICATION NUMBER: 60/196,001
PRIOR FILING DATE: 2000-04-07
NUMBER OF SEQ ID NOS: 79
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 20
LENGTH: 188
TYPE: PRT
ORGANISM: Physcomitrella patens
US-10-716-089-20

Query Match 100.0%; Score 984; DB 4; Length 188;
Best Local Similarity 100.0%; Pred. No. 6.7e-85; Indels 0; Gaps 0;
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATEVSGEITSGAPEGPMCKNLGFGSQATMGLCSKCYRETYMQAKMTALAEQATQA 60
DB 1 MATEVSGEITSGAPEGPMCKNLGFGSQATMGLCSKCYRETYMQAKMTALAEQATQA 60
QY 61 AQTATTAANVPAPVHETKLTCEVERTMIVPHOSSSYQODLVTPAAAAPQAVKSSIAA 120
DB 61 AQTATTAANVPAPVHETKLTCEVERTMIVPHOSSSYQODLVTPAAAAPQAVKSSIAA 120
QY 121 PSRPPNRCGSRKRVGLTFKRCGNYCALHRYSDKHCTYDYKKAQGEAIKAMPV 180
DB 121 PSRPPNRCGSRKRVGLTFKRCGNYCALHRYSDKHCTYDYKKAQGEAIKAMPV 180
QY 181 VAEKVKF 188
DB 181 VAEKVKF 188

RESULT 3
US-10-767-701-47138
Sequence 47138, Application US/10767701
Publication No. US20040172684A1
GENERAL INFORMATION:
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(5353)B
CURRENT APPLICATION NUMBER: US/10/767,701
CURRENT FILING DATE: 2004-01-29
NUMBER OF SEQ ID NOS: 63128
SEQ ID NO 47138
LENGTH: 171
TYPE: PRT
ORGANISM: Sorghum bicolor
FEATURE:
OTHER INFORMATION: Clone ID: SORBI-28WAY03-C3087_1.pep
US-10-767-701-47138

Query Match 43.2%; Score 425.5; DB 4; Length 171;
Best Local Similarity 46.1%; Pred. No. 3.9e-32; Indels 13; Gaps 2;
Matches 83; Conservative 28; Mismatches 56; Indels 13; Gaps 2;
QY 8 QETTSQABEGPMCKNLGFGSQATMGLCSKCYRETYMQAKMTALAEQATQAQATSAT 67
DB 4 KETGQAPBGPILCLINNGFPGSAATMNMCKCHKEMIMKQOAKLA--ASSIDSIYNGN 61
QY 68 AAAYPPAPVHETKLTCEVERTMIVPHOSSSYQODLVTPAAAAPQAVKSSIAAPSRPEPN 127
DB 62 DAVMEP-----VAGNTVVAADIEIQTMMVQPADVAGPSEGAAVISKGVGPN 110

QY 128 RCGSRKRVGLTFKRCGNYCALHRYSDKHCTYDYKKAQGEAIKAA 187
DB 111 RCGSRKRVGLTFKRCGNYCALHRYSDKHCTYDYKKAQGEAIKAA 170

RESULT 4
US-10-437-963-123596
Sequence 123596, Application US/10437963
Publication No. US20040123343A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Mc
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improc
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 123596
LENGTH: 171
TYPE: PRT
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_26415C.1.pep
US-10-437-963-123596

Query Match 42.6%; Score 419.5; DB 4; Length 1
Best Local Similarity 45.6%; Pred. No. 1.5e-31; Indels 2;
Matches 82; Conservative 26; Mismatches 59; Indels 2;

QY 8 QETTSQABEGPMCKNLGFGSQATMGLCSKCYRETYMQAKMTALAEQ 67
DB 4 KETGQAPBGPILCLINNGFPGSAATMNMCKCHKEMIMKQOAKLA-- 61
QY 68 AAAYPPAPVHETKLTCEVERTMIVPHOSSSYQODLVTPAAAAPQAVK 127
DB 62 DSGKEPIIAGHAEVAAQVEVKTIV-----AQPAEIAQPSG 110
QY 128 RCGSRKRVGLTFKRCGNYCALHRYSDKHCTYDYKKAQGEAIKAA 187
DB 111 RCGSRKRVGLTFKRCGNYCALHRYSDKHCTYDYKKAQGEAIKAA 170

RESULT 5
US-10-437-963-140292
Sequence 140292, Application US/10437963
Publication No. US20040123343A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Mc
TITLE OF INVENTION: Plants and Uses Thereof for Plant Impro
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 140292
LENGTH: 173
TYPE: PRT
ORGANISM: Oryza sativa

[illegible]

RESULT 6
 US-10-437-963-140291
 ; Sequence 140291, Application US/10437963
 ; Publication NO. US20040123343A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Wu, Wei
 ; APPLICANT: Boukharov, Andrey A.
 ; APPLICANT: Bardazuk, Brad
 ; APPLICANT: Li, Ping
 ; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53221) B
 ; CURRENT APPLICATION NUMBER: US/10/437, 963
 ; CURRENT FILING DATE: 2003-05-14
 ; NUMBER OF SEQ ID NOS: 204966
 ; SEQ ID NO 140291
 ; LENGTH: 338
 ; TYPE: PRT
 ; ORGANISM: *Oryza sativa*
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT4530_41502C.1.pep
 ; US-10-437-963-140291

Query Match	42.3%	Score 416.5	DB 4	Length 338
Bseq Local Similarity	42.7%	Pred. No. 6.3e-31		
Matches	82	Conservative	28	Matches 47; Indels 35; Gaps 3
QY	8	QETTSQAEEGPVMCGNLGCFPGSGAQTWGLGCSKCYRETVNQAKMTALAEQATQA-----	60	
Db	169	KEAGCQOQEGEPIILCTTNGCGFPGSAATMMCSCKCHEMIMKEQALTAASSIDSYNGCDG	228	
QY	61	-----AQTSTATAAIVOPPAFVHEFKLTCEVERITMIVPHQSSYYOODLTYPAAAPQAVK	115	
Db	229	GKEHIVAAAGSGSTAAVV-----AOYEAKTLV-----VQFPDVAGTSEE	265	
QY	116	SSIAAPSRPRPRCGSCSKRIVGLTGFKRCRCGNLYCALHRYSDPKHTCTYDYKAAAGEATAK	175	
Db	266	VAAPVKVVEGGRNRCATCRKRVGLTGFNCRCGMTCALHRYSDKHECPDYRTAARDATAK	325	
QY	176	ANPLVVAEKVVK	187	
Db	326	ANPVVKAEKLDK	337	

```

RESULT 7
US-10-425-115-261492
; Sequence 261492, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecu-
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425, 115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 261492
; LENGTH: 193
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_170095C.1.pep
US-10-425-115-261492

Query Match      42.2%; Score 415; DB 4; Length 19;
Best Local Similarity 47.0%; Pred. No. 4.5e-31;
Matches      85; Conservative 27; Mismatches 57; Indels 4;

```

OY	8 QETTSQAPBPVPMCKNLGFFGSQATMGLSKCYRETVMAQKATALAE	67
Dd	23 KEACCOAPEGPILTCINCGFFGSAAITNMCSKHKEMITTKODAKLA--	80
OY	68 AAAVQPAPVHETKLTCVEERTMIVPQSSSYOODLVTPA-AAAPQAVK	126
Dd	81 DAVNEPVVAGSNLT-----VVAAVAQVELQTNAVQD----PADVAGPSGV	131
OY	127 NRCSSCRKRVLGTGFKCRCGNLYCALHRYSDDKTCTTYDYAAQGEALAK	186
Dd	132 NRCSACRKRVLGTGFNCRCGNLYCALHRYSDKDCKEDYTAAARDALAY	191
OY	187 K 187	
Dd	192 K 192	

```

RESULT 8
US-10-425-115-256127
; Sequence 256127, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecu:
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
CURRENT APPLICATION NUMBER: US/10/425.115
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 369326
SEQ ID NO 256127
LENGTH: 211
TYPE: PRT
ORGANISM: Zea mays
FEATURES:
; OTHER INFORMATION: Clone ID: MRT4577_16517C.1 pep
US-10-425-115-256127

Query Match      42.2% Score 415, DB 4; Length 211
Best Local Similarity 47.0%; Pred. No. 5e-31;
Matches    85; Conservative   27; Mismatches    57; Indels    4;

cy      8 QETTQAEPGPNCKNLGFFGSQATWGLCSKCXYRVTVMQAKMTALABC
:|||||::|||::|:

```

```
Db 41 KEAGCAEBGPILICINNCGFFGSAATNMCSKCHEMITKQOAKLA--ASSIDSVNGS 98
Qy 68 AAAVQPPAPVHETKLTCEVERTMTVPHOSSSYQODLVTPA-AAAPQAVKSSIAAPSRPEP 126
Db 99 DAVMEPPVAGSNT-----VVAVAQVELQTMNVQO----PADVAGSEGVAAISKGGKVG 149
Qy 127 NRCSGCRKRVGLTGFKRCGNLYCALHRYSDKHTCTYDYKAAQOAIKAPLVVAEKV 186
Db 150 NRCSACRKRKRVGLTGFKRCGNLYCALHRYSDKHDCKFDYRTAARDALIAKANPVVADKLD 209
Qy 187 K 187
Db 210 K 210

RESULT 9
US-10-425-115-261494
; Sequence 261494, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 261494
; LENGTH: 224
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_170097C.1.pdp
US-10-425-115-261494

Query Match
Best Local Similarity 47.0%; Score 415; DB 4; Length 224;
Matches 85; Conservative 27; Mismatches 57; Indels 12; Gaps 4;

Qy 8 QETTSQABEGPVMCKNLGFFGSAATMGLCSKCYRETYWQAATMALAQAQATAT 67
Db 39 KEAGCAEBGPILICINNCGFFGSAATNMCSKCHEMITKQOAKLA--ASSIDSVNGS 96
Qy 68 AAAVQPPAPVHETKLTCEVERTMTVPHOSSSYQODLVTPA-AAAPQAVKSSIAAPSRPEP 126
Db 97 DAVMEPPVAGSNT-----VVAVAQVELQTMNVQO----PADVAGSEGVAAISKGGKVG 147
Qy 127 NRCSGCRKRVGLTGFKRCGNLYCALHRYSDKHTCTYDYKAAQOAIKAPLVVAEKV 186
Db 148 NRCSACRKRKRVGLTGFKRCGNLYCALHRYSDKHDCKFDYRTAARDALIAKANPVVADKLD 207
Qy 187 K 187
Db 208 K 208

RESULT 10
US-10-425-115-256094
; Sequence 256094, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
```

```
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 256094
; LENGTH: 171
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(171)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_16514C.1.pdp
US-10-425-115-256094

Query Match
Best Local Similarity 43.1%; Score 414.5; DB 4; Length 171
Matches 84; Conservative 26; Mismatches 42; Indels 4;

Qy 8 QETTSQABEGPVMCKNLGFFGSAATMGLCSKCYRETYWQAATMALAEC 67
Db 4 KETGCAEBGPILICINNCGFFGSAATNMCSKCHEMIMKQEOAKLA-- 61
Qy 68 AAAVQPPAPVHETKLTCEVERTMTVPHOSSSYQODLVTPA-AAAPQAVK 121
Db 62 DAVMEPPVAGSNT-----VVAVAQVELQTMNVQO----PADVAGSEGVAAISKGGKVG 147
Qy 122 S-----RPEPNRCSGCRKRVGLTGFKRCGNLYCALHRYSDKHTC 172
Db 96 SXGGGGLQREGXPNRCSGCRKRVGLTGFKRCGNLYCALHRYSDKHTC 155
Qy 173 IAKANPLVVAEKV 187
Db 156 IAKANPVVAKDKDK 170

RESULT 11
US-10-425-114-59729
; Sequence 59729, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules
; FILE REFERENCE: 38-21(53133)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 59729
; LENGTH: 179
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3912-048-B5_FLI.pdp
US-10-425-114-59729

Query Match
Best Local Similarity 41.9%; Score 412.5; DB 4; Length 171
Matches 80; Conservative 28; Mismatches 50; Indels 3;

Qy 8 QETTSQABEGPVMCKNLGFFGSAATMGLCSKCYRETYWQAATMALAEC 60
Db 12 KEAGCAEBGPILICINNCGFFGSAATNMCSKCHEMIMKQEOAKLA-- 71
Qy 61 --AQTSAATAAVQPPAPVHETKLTCEVERTMTVPHOSSSYQODLVTPA 118
Db 72 GKGRALIAITGVAVP-----QVEKTTA-----VQPM 109
Qy 119 AAPSRPEPNRCSGCRKRVGLTGFKRCGNLYCALHRYSDKHTCTYDYKA 178
```


Db 110 IAAKGGPNRCATCRKRVLTGFCNRCGNTYCSMRYSDDKHQCFDYRTAARDALAKNP 169

QY 179 LVVAEKRVK 187
:|||:|

Db 170 VVAEKLDK 178

```

RESULT 12
US-10-425-114-65423
/ Sequence 65423, Application US/10425114
/ Publication No. US2004003488A1
/ GENERAL INFORMATION:
/ APPLICANT: Liu, Jingdong
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Screen, Steven E
/ APPLICANT: Tadaska, Jack E
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
/ FILE REFERENCE: 38-21(5313)B
/ CURRENT APPLICATION NUMBER: US/10/425,114
/ NUMBER OF SEQ ID NOS: 2003-04-28
/ SEQ ID NO 65423
/ LENGTH: 179
/ TYPE: PRT
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: LIB4764-003-Fl1_Fl1.pep
/ US-10-425-114-65423

```

Query Match	41.9%	Score 412.5;	DB 4;	Length 179;
Best Local Similarity	42.3%	Pred. No. 7e-31;		
Matches 80; Conservative	28;	Mismatches 50;	Indels 31;	Gaps 3;

```

QY      8 QEETSAPEGPMWCKNLCGFFSQATMGICSKCRETMAKMTALAEQATQA----- 60
Db      12 KKAEGCQPEPGPILLCINNCGFSSAAITNNCKSKCHEKIMIKOEQAQLAASIDSTIVNGDN 71
QY      61 --AQSATATAAVQPPAPVHEHTKLTCSEVERTMIVPHOSSYQODLVTPAALAPQAVKSI 118
Db      72 GKGPAALATVGVAVP-----QVEBKTLA-----VQPMVALETSEAAV 109
QY      119 AAPSPREPNNCCSKRNUGLTGFKRCGNLYCALHRSDDKTCITYDKAAGEALAKAFN 178
Db      110 IAKAKGPMRCATCRKRVGLTGFNCRGNCTYCSMHRSYDKDKDCQFDVRYTAADALAKAFN 169
QY      179 LVVAEKVKK 187
Db      170 VVKAERGLDK 178

```

```

RESULT 13
US-10-767-701-45710
; Sequence 45710, Application US/10767701
; Publication NO. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ. ID NOS: 63128
; SEQ. ID NO 45710
; LENGTH: 171
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MA03-C6463.1 pep

```

US-10-767-701-45710

Query Match	41.8%;	Score 411.5;	DB 4;	Length :
Best Local Similarity	43.2%;	Pred. No. 8.3e-31;		
Matches	83;	Conservative	27;	Mismatches 45;
				Indels

[illegible]

```

RESULT 14
US-10-437-963-123595
Sequence 123595, Application US/10437963
Publication No. US2004012343A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53221) B
CURRENT APPLICATION NUMBER: US/10/437, 963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 123595
LENGTH: 171
TYPE: PRT
ORGANISM: Oryza sativa
FEATURES:
OTHER INFORMATION: Clone ID: PAT_MRT4530_26414C.1 pep
US-10-437-963-123595

```

Query Match	41.4%	Score 407.5	DB 4	Length 1
Best Local Similarity	45.0%	Pred. No. 2e-30		
Matches 81	Conservative 25	Mismatches 61	Indels 2	

[illegible]

RESULT 15
US-10-425-115-308732
; Sequence 308732, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 2005, 16:29:16 ; Search time 25 Seconds
(without alignments)
41.995 Million cell updates/sec

Title: US-10-716-089-20

Perfect score: 984
Sequence: 1 WATERVSOEITSGAPEGPM.....GOEAIKANPLVVAEKVKF 188

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 32527 seqs, 5584426 residues

Total number of hits satisfying chosen parameters: 32527

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA New:*
1: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB pep.*
2: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB pep.*
3: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB pep.*
4: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB pep.*
5: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB pep.*
6: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB pep.*
7: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB pep.*
8: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	88	8.9	595	US-11-182-946-9	Sequence 9, Appl1
2	80	8.1	2432	US-10-821-234-899	Sequence 899, App
3	79	8.0	259	US-10-131-826A-300	Sequence 300, App
4	79	8.0	259	US-11-182-946-2	Sequence 2, Appl1
5	76.5	7.8	1163	US-11-044-899-2	Sequence 2, Appl1
6	76.5	7.8	1163	US-11-044-899-30	Sequence 30, Appl1
7	72.5	7.4	190	US-11-055-822-538	Sequence 538, App
8	72	7.3	2630	US-11-186-731-2	Sequence 2, Appl1
9	72	7.3	7968	US-11-186-731-5	Sequence 5, Appl1
10	71.5	7.3	502	US-11-021-441-7	Sequence 7, Appl1
11	71.5	7.3	563	US-11-021-441-9	Sequence 9, Appl1
12	71.5	7.3	574	US-11-021-441-15	Sequence 15, Appl1
13	71.5	7.3	581	US-11-021-441-11	Sequence 11, Appl1
14	71.5	7.3	581	US-11-021-441-13	Sequence 13, Appl1
15	71.5	7.3	1035	US-11-021-441-4	Sequence 4, Appl1
16	71	7.2	715	US-10-131-826A-116	Sequence 116, App
17	70.5	7.2	1207	US-10-821-234-1105	Sequence 1105, App
18	70.5	7.2	4419	US-10-821-234-1155	Sequence 1155, App
19	69.5	7.1	1066	US-11-055-822-370	Sequence 370, App
20	69.5	7.1	1066	US-11-055-822-1002	Sequence 1002, App
21	69.5	7.1	1113	US-11-055-822-368	Sequence 368, App
22	69.5	7.1	1113	US-11-055-822-1000	Sequence 1000, App
23	69	7.0	188	US-11-064-774A-117	Sequence 117, App
24	69	7.0	1377	US-10-821-234-1070	Sequence 1070, App
25	68.5	7.0	1005	US-11-113-424-63	Sequence 63, Appl1

26	68	6.9	151	6	US-10-821-234-1378	Sec	App
27	68	6.9	151	6	US-10-528-031-9	Sec	App1
28	68	6.9	409	6	US-10-878-556A-55	Sec	App1
29	68	6.9	943	6	US-10-467-657-5508	Sec	App
30	67.5	6.9	481	6	US-10-467-657-4396	Sec	App
31	67	6.8	985	7	US-11-113-424-61	Sec	App1
32	66.5	6.8	463	6	US-10-131-826A-160	Sec	App
33	66.5	6.8	3507	7	US-11-075-185-7	Sec	App
34	66	6.7	251	7	US-11-054-515-1538	Sec	App
35	66	6.7	442	6	US-10-821-234-1594	Sec	App
36	66	6.7	712	6	US-10-770-726-66	Sec	App1
37	66	6.7	1178	7	US-11-044-899-29	Sec	App1
38	65	6.6	346	6	US-10-131-826A-432	Sec	App
39	65	6.6	346	6	US-10-967-457-77	Sec	App1
40	65	6.6	358	6	US-10-821-234-878	Sec	App
41	64.5	6.6	1199	6	US-10-821-234-1126	Sec	App
42	64	6.5	465	6	US-10-821-234-1550	Sec	App
43	63.5	6.5	550	6	US-10-878-556A-166	Sec	App
44	63.5	6.5	669	6	US-10-878-556A-87	Sec	App1
45	63	6.4	219	6	US-10-689-742-106	Sec	App

ALIGNMENTS

RESULT 1
US-11-182-946-9
; Sequence 9, Application US/11182946
; Publication No. US20050255100A1
; GENERAL INFORMATION:
; APPLICANT: Wei, Ying-Fei
; APPLICANT: Ni, Jian
; APPLICANT: Gentz, Reiner
; APPLICANT: Ruben, Steven
; TITLE OF INVENTION: Tumor Necrosis Factor Receptor 5
; FILE REFERENCE: 1488.1280004
; CURRENT APPLICATION NUMBER: US/11/182,946
; CURRENT FILING DATE: 2005-07-18
; PRIOR APPLICATION NUMBER: US/10/186,643
; PRIOR FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: US/09/573,986
; PRIOR FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 9
; LENGTH: 595
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-182-946-9

Query Match 8.9%; Score 88; DB 7; Length 595.
Best Local Similarity 18.5%; Pred. No. 0.13;
Matches 48; Conservative 30; Mismatches 98; Indels 8;

QY	11	TSQAEPEVMCKNLC-----	29
DB	55	TQQCGQRPPTCKQCEPDYUDEADRCTACTYCSRDILVEKTPCAMNSI	114
QY	30	SOATGGLCSKCYRETV-----MOAKMTALAEQATQAAQTSATPAAVQPI	83
DB	115	STSAVNSCARCFHSVYCAAGMIVKFPGRNOKTYCEPSPVSPACASFI	174
QY	84	CEVERTMIVPHOSS-----YQDDLVTPAAAPQAVKSIADSE	125
DB	175	FOAKTPTVPSPATSSASTVPVRGTRTLAGAASKLTRAADS-PSSEVGRPS	233
QY	126	ENRQSCSKRGVLGTFKRCGN-----LYCALHRSYDKHCTCYD-----	167
DB	234	PEGSDCKKQCEPDYUDEADRCTACTYCSRDILVEKTPCAMNSRTCE	293
QY	168	AGOEAIKANPLVVAEKVK 187	
DB	294	TNSCARCVPIPCAAETVTK 313	

```
RESULT 2
US-10-821-234-899
; Sequence 899, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crahn, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; PRIOR FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pc_seq_genes Version 1.0
; SEQ ID NO 899
; LENGTH: 2432
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-899

Query Match      8.1%; Score 80; DB 6; Length 2432;
Best Local Similarity 21.8%; Pred. No. 3.7;
Matches 37; Conservative 37; Mismatches 78; Indels 18; Gaps 6;

QY 5 RVSETTQAPEGPVWCKNLGPFQSQATMGLCSKCYRETVM---QAKWTALAEQATPA 61
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 984 RIAPRPYLAPR-PLMLASRSMWMSYAERSMMSYERSMMSYERSMMSYPAERSMMSA 1042
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 62 QATSAATAAQP-PAVPHETKLTCEVERTWIVPHOSSSYQDVLTPAA-----APQA 113
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1043 YERSMMSYERSMMSYPAERSMMSYERSMMSYERSMMSYERSMMSYPAERSMMSA 1097
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 114 VVSSIAABSRPEPNCGSCRKRVGLTFPCRCGNLYCLHRSDDHCTCY 163
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1098 MMSYSAADRSMSST-SAADRSMSSTYADRSMSMADSDYDSDYD 1146
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 3
US-10-131-826A-300
; Sequence 300, Application US/10131826A
; Publication No. US20050245730A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerltsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C128
; CURRENT APPLICATION NUMBER: US/10/131,826A
; CURRENT FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
```

```

; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; Remaining Prior Application data removed - See file wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 300
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-131-826A-300

Query Match      8.0%; Score 79; DB 6; Length 259;
Best Local Similarity 22.2%; Pred. No. 0.38;
Matches 34; Conservative 19; Mismatches 62; Indels 6;

QY 6 VSETTQAPEG-----PVMCKNLGPFQSQATMGLCS-----KCYR 52
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 103 MTRDTVCQCKEGTFRNENSPMCKRCKSCPCGEGVQVNSCTSWDDIQCV 162
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 53 LAEQATQAQATSATAA-----AVQPPAPVHEKLTCE-----VVR 99
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 163 AAEETMNTSPGTPAPAAEETMNTSPGTPAPAAEETMNTSPGTPAPAAE 213
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 100 QQDLVTPAAPAAPOAVKSSIAAPSRPEPNRCGSC 132
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 214 ---TTSPTGPAPAAEETMTTSPGTPTASHTYLS 243
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 4
US-11-182-946-2
; Sequence 2, Application US/11182946
; Publication No. US20050255100A1
; GENERAL INFORMATION:
; APPLICANT: Wei, Ying-Fei
; APPLICANT: Ni, Jian
; APPLICANT: Gentz, Reiner
; APPLICANT: Ruben, Steven
; TITLE OF INVENTION: Tumor Necrosis Factor Receptor 5
; FILE REFERENCE: 1488.1280004
; CURRENT APPLICATION NUMBER: US/11/182,946
; CURRENT FILING DATE: 2005-07-18
; PRIOR APPLICATION NUMBER: US/10/186,643
; PRIOR FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: US/09/573,986
; PRIOR FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-182-946-2

Query Match      8.0%; Score 79; DB 7; Length 259;
Best Local Similarity 22.2%; Pred. No. 0.38;
Matches 34; Conservative 19; Mismatches 62; Indels 6;

QY 6 VSETTQAPEG-----PVMCKNLGPFQSQATMGLCS-----KCYR 52
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 103 MTRDTVCQCKEGTFRNENSPMCKRCKSCPCGEGVQVNSCTSWDDIQCV 162
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```



```

Oy      28 FGSQATWGL-----CSKCYRE--TWQAKMTALAEQATQATSTPAAVGPAPVHE 79
Db      8 FGLIATTAALSAALVACSPHPHQDDSPVQGRATNEILITTSQNPITSASSITSATTSAPVEE 67

Oy      80 TKLTCEVERTWIVPHQSSSYQOODVTPPAAAPOAVKSSIAAPSPREPNCRG 130
Db      68 -DVEIIVSPPALVDEQVTFEISGLDEPGGYAAICDSVANPGRVPVPSCTG 117

RESULT 8
US-11-186-731--2
; Sequence 2, Application US/11186731
; Publication No. US20050255521A1
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Acton, Susan L.
; TITLE OF INVENTION: 59079 and 12599, Protein Kinase Family
; TITLE OF INVENTION: Members and Uses Therefor
; FILE REFERENCE: MPI2001-047PRC1 (M)
; CURRENT APPLICATION NUMBER: US/11/186,731
; CURRENT FILING DATE: 2005-07-21
; PRIOR APPLICATION NUMBER: US/10/077,130
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: 60/269201
; PRIOR FILING DATE: 2001-02-15
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 2630
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-186-731-2

```

	Query Match	7.3%	Score 72;	DB 7;	Length 2630;
	Beet Local Similarity	26.5%;	Pred. No. 24;		
	Matches	Conservative 15;	Indels 32;	Gaps 5;	
Oy	24 LCGFFGSANTGTCSCYRETVMAQKMTALAEQAQAATSATATAAQVAPAPVHETLTT	:	:	:	83
Dd	1564 LKGVIYAGALPGL-----REPMEHRV--LEEARREQAAT-----LLAKRPFETALR	:	:	:	1610
Oy	84 CEVERTMIVPQSSSYOODLTVPAAAAQAVKSSIARSPDEPN--RCSGCRK	:	:	:	134
Dd	1611 LPASGTTHLPAGHSLSHD-----SPTSPPSSEACGEAQR	:	:	:	1646

```

RESULT 9
US-11-186-731-5
; Sequence 5, Application US/11186731
; Publication No. US20050255521A1
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Accon, Susan L.
; TITLE OF INVENTION: 59079 and 12599, Protein Kinase Family
; TITLE OF INVENTION: Members and Uses Therefor
; FILE REFERENCE: MP12001-047P1RCP1(W)
; CURRENT APPLICATION NUMBER: US/11/186, 731
; CURRENT FILING DATE: 2005-07-21
; PRIOR APPLICATION NUMBER: US/10/077,130
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: 60/269201
; PRIOR FILING DATE: 2001-02-15
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 7968
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-186-731-5

Query Match      7.3%   Score 72, DB 7, Length 7968;
Best Local Similarity 26.5%;   Pred. No. 81;

```

	Matches	30; Conservative	15; Mismatches	36; Indels	5;
Oy	24	LCGFSGQATMGLCKCYRHWQAKRTALAEQTAAQTSAATAAAC			83
Db	6302	LKGGITVAGLPGL-----REPLSEHRV--LEENAREQAT-----IL			6948
Oy	84	CEVERTMIVPHQSSSYQODLVTPAAAPPAQVKSIAAPSRPEPV--RCG			
Db	6949	LPASGTHIALPGHSHSLHD-----SPSTRPSSEACG			

```

RESULT 10
US-11-021-441-7
/ Sequence 7, Application US/11021441
/ Publication No. US20050245748A1
/ GENERAL INFORMATION:
/ APPLICANT: DUBENSKY, Thomas W., Jr.
/ APPLICANT: PORTNOY, Daniel A.
/ APPLICANT: LUCKETT, William S., Jr.
/ APPLICANT: COOK, David N.
/ TITLE OF INVENTION: RECOMBINANT NUCLEIC ACID MOLECULES,
/ TITLE OF INVENTION: EXPRESSION CASSETTES, AND BACTERIA, AND
/ TITLE OF INVENTION: THEREOF
/ FILE REFERENCE: 282172003900
/ CURRENT APPLICATION NUMBER: US/11/021,441
/ CURRENT FILING DATE: 2004-12-23
/ PRIOR APPLICATION NUMBER: US 60/616,750
/ PRIOR FILING DATE: 2004-10-06
/ PRIOR APPLICATION NUMBER: US 60/615,287
/ PRIOR FILING DATE: 2004-10-01
/ PRIOR APPLICATION NUMBER: US 60/599,377
/ PRIOR FILING DATE: 2004-08-05
/ PRIOR APPLICATION NUMBER: PCT/US2004/23881
/ PRIOR FILING DATE: 2004-07-23
/ PRIOR APPLICATION NUMBER: US 10/883,599
/ PRIOR FILING DATE: 2004-06-30
/ PRIOR APPLICATION NUMBER: US 60/556,744
/ PRIOR FILING DATE: 2004-03-26
/ NUMBER OF SEQ ID NOS: 129
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 7
/ LENGTH: 502
/ TYPE: prt
/ ORGANISM: Homo sapiens
US-11-021-441-7

```

```

Query Match 7.3%; Score 71.5; DB 7; Length 50
Best Local Similarity 22.8%; Pred. No. 4.3;
Matches 34; Conservative 18; Mismatches 62; Indels

QY      15 PEGPYMC-----KNLC-----GFGSQATWGLCKSKCYRETV--MQAKU      60
           |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB      212 PIGGLCOAGYKEVDACQACSPGPFKEABESPCLECPHTLLPSEGA'      291
           |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

QY      61 AQAISATLAAVQPPAPVH---ETKLTCVESTMLVPHSSSYQODLVTF'      117
           |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

DB      292 APQDPASMPCTRPPEAPHYLTAVGAGCAVETLRW--TPPDQSGREDI---      340
           |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

QY      118 IAAERPEPNRCGSCRRKV-----GLT 139
           |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

DB      341 TCBOCWPESGEGCPCEASVRISEPHGLT 369

RESULT 11
US-11-021-441-9
; Sequence 9, Application US/11021441
; Publication No. US20050249748A1
; GENERAL INFORMATION:
; APPLICANT: DUBENSKY, Thomas W., Jr.
; APPLICANT: PORTNOY, Daniel A.
; APPLICANT: LUCKETT, William S., Jr.
; APPLICANT: COOK, David N.
; TITLE OF INVENTION: RECOMBINANT NUCLEIC ACID MOLECULES,

```

```

; TITLE OF INVENTION: EXPRESSION CASSETTES, AND BACTERIA, AND METHODS OF USE
; FILE OF INVENTION: THEROP
; FILE REFERENCE: 282172003900
; CURRENT APPLICATION NUMBER: US/11/021,441
; CURRENT FILING DATE: 2004-12-23
; PRIOR APPLICATION NUMBER: US 60/616,750
; PRIOR FILING DATE: 2004-10-06
; PRIOR APPLICATION NUMBER: US 60/615,287
; PRIOR FILING DATE: 2004-10-01
; PRIOR APPLICATION NUMBER: US 60/599,377
; PRIOR FILING DATE: 2004-08-05
; PRIOR APPLICATION NUMBER: PCT/US2004/23881
; PRIOR FILING DATE: 2004-07-23
; PRIOR APPLICATION NUMBER: US 10/883,599
; PRIOR FILING DATE: 2004-06-30
; PRIOR APPLICATION NUMBER: US 60/556,744
; PRIOR FILING DATE: 2004-03-26
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 563
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion protein
US-11-021-441-9

```

```

Query Match      7.3%; Score 71.5; DB 7; Length 563;
Best Local Similarity 22.8%; Pred. No. 4.8;
Matches 34; Conservative 18; Mismatches 62; Indels 35; Gaps 7;

```

```

QY 15 PEGPVMC-----KNLC-----GFGSOATMGLCSKCYRETV--MQAKMTALAEQATQA 60
DB 293 PFGQCLCGAGYKVEDACQACSPGFFKFEASSPCLCEPHTLPSPGATSCCEGEPFR 352
QY 61 AQAATATAAAGVAPPAVH---ETKLTCEVERTMIIPHOSSYQODLVTPAALAPQAVKSS 117
DB 353 APQDPAAMPCTRPPSAPHYLTAVGMGAKVELRW--TPPDGSGREDI-----VTSV 401
QY 118 IAAPSRPEPNRCGSCRKRY-----GLT 139
DB 402 TCEQCWPESGECGPCEASVRYSEBPPHGLT 430

```

```

RESULT 12
US-11-021-441-15
; Sequence 15, Application US/11021441
; Publication No. US20050249748A1
; GENERAL INFORMATION:
; APPLICANT: DUBENSKY, Thomas W., Jr.
; APPLICANT: PORTNOY, Daniel A.
; APPLICANT: LUCKETT, William S., Jr.
; APPLICANT: COOK, David N.
; TITLE OF INVENTION: RECOMBINANT NUCLEIC ACID MOLECULES,
; TITLE OF INVENTION: EXPRESSION CASSETTES, AND BACTERIA, AND METHODS OF USE
; TITLE OF INVENTION: THEROP
; FILE REFERENCE: 282172003900
; CURRENT APPLICATION NUMBER: US/11/021,441
; CURRENT FILING DATE: 2004-12-23
; PRIOR APPLICATION NUMBER: US 60/616,750
; PRIOR FILING DATE: 2004-10-06
; PRIOR APPLICATION NUMBER: US 60/615,287
; PRIOR FILING DATE: 2004-10-01
; PRIOR APPLICATION NUMBER: US 60/599,377
; PRIOR FILING DATE: 2004-08-05
; PRIOR APPLICATION NUMBER: PCT/US2004/23881
; PRIOR FILING DATE: 2004-07-23
; PRIOR APPLICATION NUMBER: US 10/883,599
; PRIOR FILING DATE: 2004-06-30
; PRIOR APPLICATION NUMBER: US 60/556,744
; PRIOR FILING DATE: 2004-03-26
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSeq for Windows Version 4.0

```

```

; SEQ ID NO 15
; LENGTH: 574
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion protein
US-11-021-441-15

```

```

Query Match      7.3%; Score 71.5; DB 7; Length 574
Best Local Similarity 22.8%; Pred. No. 4.9;
Matches 34; Conservative 18; Mismatches 62; Indels 7;

```

```

QY 15 PEGPVMC-----KNLC-----GFGSOATMGLCSKCYRETV--MQAK 60
DB 294 PFGQCLCGAGYKVEDACQACSPGFFKFEASSPCLCEPHTLPSPGATSCCEGEPFR 353
QY 61 AQAATATAAAGVAPPAVH---ETKLTCEVERTMIIPHOSSYQODLVTP 117
DB 354 APQDPAAMPCTRPPSAPHYLTAVGMGAKVELRW--TPPDGSGREDI--- 402
QY 118 IAAPSRPEPNRCGSCRKRY-----GLT 139
DB 403 TCEQCWPESGECGPCEASVRYSEBPPHGLT 431

```

```

RESULT 13
US-11-021-441-11
; Sequence 11, Application US/11021441
; Publication No. US20050249748A1
; GENERAL INFORMATION:
; APPLICANT: DUBENSKY, Thomas W., Jr.
; APPLICANT: PORTNOY, Daniel A.
; APPLICANT: LUCKETT, William S., Jr.
; APPLICANT: COOK, David N.
; TITLE OF INVENTION: RECOMBINANT NUCLEIC ACID MOLECULES,
; TITLE OF INVENTION: EXPRESSION CASSETTES, AND BACTERIA, AND
; TITLE OF INVENTION: THEROP
; FILE REFERENCE: 282172003900
; CURRENT APPLICATION NUMBER: US/11/021,441
; CURRENT FILING DATE: 2004-12-23
; PRIOR APPLICATION NUMBER: US 60/616,750
; PRIOR FILING DATE: 2004-10-06
; PRIOR APPLICATION NUMBER: US 60/615,287
; PRIOR FILING DATE: 2004-10-01
; PRIOR APPLICATION NUMBER: US 60/599,377
; PRIOR FILING DATE: 2004-08-05
; PRIOR APPLICATION NUMBER: PCT/US2004/23881
; PRIOR FILING DATE: 2004-07-23
; PRIOR APPLICATION NUMBER: US 10/883,599
; PRIOR FILING DATE: 2004-06-30
; PRIOR APPLICATION NUMBER: US 60/556,744
; PRIOR FILING DATE: 2004-03-26
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 581
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion protein
US-11-021-441-11

```

```

Query Match      7.3%; Score 71.5; DB 7; Length 581
Best Local Similarity 22.8%; Pred. No. 5;
Matches 34; Conservative 18; Mismatches 62; Indels 7;

```

```

QY 15 PEGPVMC-----KNLC-----GFGSOATMGLCSKCYRETV--MQAK 60
DB 301 PFGQCLCGAGYKVEDACQACSPGFFKFEASSPCLCEPHTLPSPGATSCCEGEPFR 360
QY 61 AQAATATAAAGVAPPAVH---ETKLTCEVERTMIIPHOSSYQODLVTP 117
DB 361 APQDPAAMPCTRPPSAPHYLTAVGMGAKVELRW--TPPDGSGREDI--- 409

```

```
OY 118 IAAPSRPEPNRCGSCRKRV-----GLT 139
DB 410 TCBCQWPESSGCGPCASVRYSEPPHGLT 438

RESULT 14
US-11-021-441-13
; Sequence 13, Application US/11021441
; Publication NO. US20050249748A1
; GENERAL INFORMATION:
; APPLICANT: DUBENSKY, Thomas W., Jr.
; APPLICANT: PORTNOY, Daniel A.
; APPLICANT: LUCKETT, William S., Jr.
; APPLICANT: COOK, David N.
; TITLE OF INVENTION: RECOMBINANT NUCLEIC ACID MOLECULES,
; TITLE OF INVENTION: EXPRESSION CASSETTES, AND BACTERIA, AND METHODS OF USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 282172003900
; CURRENT APPLICATION NUMBER: US/11/021,441
; PRIOR FILING DATE: 2004-12-23
; PRIOR APPLICATION NUMBER: US 60/616,750
; PRIOR FILING DATE: 2004-10-06
; PRIOR APPLICATION NUMBER: US 60/615,287
; PRIOR FILING DATE: 2004-10-01
; PRIOR APPLICATION NUMBER: US 60/599,377
; PRIOR FILING DATE: 2004-08-05
; PRIOR APPLICATION NUMBER: PCT/US2004/23881
; PRIOR FILING DATE: 2004-07-23
; PRIOR APPLICATION NUMBER: US 10/883,599
; PRIOR FILING DATE: 2004-06-30
; PRIOR APPLICATION NUMBER: US 60/556,744
; PRIOR FILING DATE: 2004-03-26
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 581
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion protein
; US-11-021-441-13

Query Match 7.3%; Score 71.5; DB 7; Length 581;
Best Local Similarity 22.8%; Pred. No. 5;
Matches 34; Conservative 18; Mismatches 62; Indels 35; Gaps 7;

OY 15 PEGPVMC-----KNLC-----GFGSQATWGLCSKCYRETV--MQAKTALAEQATQA 60
DB 301 PIGQCLCOAGYEKVEDACQACSPGFKFEASBPCLCEPHTLPSPEGATSCCEGEGFR 360
OY 61 AQTATATAAQQPAPVH---ETKLTCVEKRTMIVPHOSSYQODLVTPAAAAQAVKSS 117
DB 361 APQDPAHPCTRPSPAPHYLTAVGMGAKVELRW--TPQDSGGREDI-----VYSV 409
OY 118 IAAPSRPEPNRCGSCRKRV-----GLT 139
DB 410 TCBCQWPESSGCGPCASVRYSEPPHGLT 438

RESULT 15
US-11-021-441-4
; Sequence 4, Application US/11021441
; Publication NO. US20050249748A1
; GENERAL INFORMATION:
; APPLICANT: DUBENSKY, Thomas W., Jr.
; APPLICANT: PORTNOY, Daniel A.
; APPLICANT: LUCKETT, William S., Jr.
; APPLICANT: COOK, David N.
; TITLE OF INVENTION: RECOMBINANT NUCLEIC ACID MOLECULES,
; TITLE OF INVENTION: EXPRESSION CASSETTES, AND BACTERIA, AND METHODS OF USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 282172003900
```

```
; CURRENT APPLICATION NUMBER: US/11/021,441
; CURRENT FILING DATE: 2004-12-23
; PRIOR APPLICATION NUMBER: US 60/616,750
; PRIOR FILING DATE: 2004-10-06
; PRIOR APPLICATION NUMBER: US 60/615,287
; PRIOR FILING DATE: 2004-10-01
; PRIOR APPLICATION NUMBER: US 60/599,377
; PRIOR FILING DATE: 2004-08-05
; PRIOR APPLICATION NUMBER: PCT/US2004/23881
; PRIOR FILING DATE: 2004-07-23
; PRIOR APPLICATION NUMBER: US 10/883,599
; PRIOR FILING DATE: 2004-06-30
; PRIOR APPLICATION NUMBER: US 60/556,744
; PRIOR FILING DATE: 2004-03-26
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1035
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion protein
; US-11-021-441-4

Query Match 7.3%; Score 71.5; DB 7; Length 10
Best Local Similarity 22.8%; Pred. No. 9.5;
Matches 34; Conservative 18; Mismatches 62; Indels 7;

OY 15 PEGPVMC-----KNLC-----GFGSQATWGLCSKCYRETV--MQAK 60
DB 315 PIGQCLCOAGYEKVEDACQACSPGFKFEASBPCLCEPHTLPSPEGA 374
OY 61 AQTATATAAQQPAPVH---ETKLTCVEKRTMIVPHOSSYQODLVTF 117
DB 375 APQDPAHPCTRPSPAPHYLTAVGMGAKVELRW--TPQDSGGREDI--- 423
OY 118 IAAPSRPEPNRCGSCRKRV-----GLT 139
DB 424 TCBCQWPESSGCGPCASVRYSEPPHGLT 452
```

Search completed: December 8, 2005, 16:36:29
Job time : 25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: December 8, 2005, 16:33:01 ; Search time 3238 Seconds
(without alignment)
3300.361 Million cell updates/sec

Title: US-10-716-089-20

Perfect score: 984

Sequence: 1 MATERVSETTSQAREGPVM.....GOEALAKNPVAKKVKVF 188

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame.p2n.model -DEV=xlp
-O=/cgn2.1/USPTO.spool.p/US10716089/cunat.05122005.094343.13141/app.query.fasta_1.327
-DB=GenEmbl -QFMT=fastap -SUFFIX=sp2n.rge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human0.cdi -LIST=45
-DOCALLIGN=200 -THR.SCORE=Pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=Pct -NORM=exc -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10716089@cgn.1.1.7415@cunat.05122005.094343.13141 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG.SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:.*
1: gb_ba:.*
2: gb_in:.*
3: gb_env:.*
4: gb_com:.*
5: gb_ov:.*
6: gb_pal:.*
7: gb_ph:.*
8: gb_pr:.*
9: gb_ro:.*
10: gb_srs:.*
11: gb_sy:.*
12: gb_un:.*
13: gb_vl:.*
14: gb_hcg:.*
15: gb_pl:.*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	984	100.0	804	6	AR452950 Sequence
2	984	100.0	804	6	AX288141 Sequence
3	957	97.3	1531	6	AR452942 Sequence

4	957	97.3	1531	6	AX288133	AX28	2e
5	419.5	42.6	999	15	AK104562	AK1	sat
6	419.5	42.6	1000	15	AK104175	AK3	sat
7	419.5	42.6	1009	15	AK104605	AK:	sat
8	419.5	42.6	1021	15	AK060008	AK0	sat
9	419.5	42.6	1209	15	AK069720	AK0	sat
10	419.5	42.6	110000	15	AP008212_240	AK0	sat
11	419.5	42.6	110000	15	AP008212_241	AK0	sat
12	419.5	42.6	158321	15	AP003626	AK0	sat
13	419.5	42.6	163568	15	AP003711	AK0	sat
14	416.5	42.3	522	6	AK654090	AX6:	2e
15	416.5	42.3	991	15	AK061911	AK0	sat
16	416.5	42.3	1000	15	AK067403	AK0	sat
17	416.5	42.3	110000	15	AP008208_033	AK0	sat
18	416.5	42.3	142732	15	AP004184	AK0	sat
19	415.5	42.2	1132	15	AY345599	AK3	sat
20	415	42.2	1020	15	BT016379	BT0	sat
21	413	42.0	1162	15	AY515607	AY5	sat
22	409.5	41.6	1026	15	BT014337	BT0	sat
23	407.5	41.4	117676	15	AP006427	AP0	sat
24	400.5	40.7	112882	14	AC161406	AC1	sat
25	390.5	39.7	100208	15	AC155881	AC1	sat
26	387.5	39.4	1032	15	BT001984	BT0	sat
27	387	39.3	941	15	BT017601	BT0	sat
28	386	39.2	798	6	AK653271	AX6:	2e
29	386	39.2	854	15	AY377427	AY3	sat
30	386	39.2	110000	15	AP008209_327	AK0	sat
31	386	39.2	162497	15	AC090871	AC0	sat
32	385.5	39.2	125195	15	AC154034	AC:	190
33	385	39.1	1158	15	AK102503	AK:	sat
34	384.5	39.1	486	6	AK654221	AX6:	2e
35	384.5	39.1	727	15	AY287740	AY:	sat
36	384.5	39.1	887	15	AK121813	AK:	sat
37	384.5	39.1	110000	15	AP008213_036	AK0	sat
38	384.5	39.1	110000	15	AP008213_037	AK0	sat
39	384.5	39.1	151986	15	AP005840	AP0	sat
40	383.5	38.0	102653	14	AY056440	AY0	sat
41	382.5	38.9	898	15	AY056445	AY0	sat
42	379	38.5	522	15	BT015115	BT0	sat
43	379	38.5	105807	15	AC006085	AC0	sat
44	371	37.7	2786	5	CR761102	CR7:	3 e
45	369.5	37.6	2161	5	BC081266	BC08	1

ALIGNMENTS

RESULT 1
AR452950
LOCUS AR452950
DEFINITION Sequence 12 from patent US 6677504.
ACCESSION AR452950
VERSION AR452950.1 GI:42685097
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
Patent: US 6677504-A 12 13-JAN-2004;
BASF Plant Science GmbH; Ludwigshafen;
DEX;

FEATURES
source
location/Qualifiers
1..804
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Alignment Scores:
Pred. No.: 6,32e-61 Length: 804
Score: 984.00 Matches: 188
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0
US-10-716-089-20 (1-188) x AR452950 (1-804)			
OY	1	MetAlaThrGluArgValSerGlnGluThrThrSerGlnAlaProGluGlyProValMet	20
DB	134	ATGGCCACCGAGCGTGTCTCAGAGACGACCTGGCAGGCCCTCGAGGGTCCAGTTATG	193
OY	21	CysLysAsnLeuCyseGlyPhePheGlySerGlnAlaThrMetGlyLeuCyseSerLysCyS	40
DB	194	TGCAAGAACCTTTGCGGCTTCTTCGGCAGCCAACTGAGGGTGTGCTCGAAGTGC	253
OY	41	TyrArgGluThrValMetGlnAlaLysMetThrAlaLeuAlaGluGlnAlaThrGlnAla	60
DB	254	TACCGAGACATGATGACGAGCAAGATGACGGCTTTAGCTGACCAAGCCCTAGGCT	313
OY	61	AlaGlnAlaThrSerAlaThrAlaAlaValAlaGlnProProAlaProValHisGluThr	80
DB	314	GCTCAGGCGACATCTGCCACAGCTGCTGTTTCAGCCCCCTCCTGTCATGAGACC	373
OY	81	LysLeuThrCyseGluValGluArgThrMetIleValProHisGlnSerSerSerTyrGln	100
DB	374	AAGCTCACATGCGAGGTGAGAGAACATGATTGTGCCATCAATCTCCAGCTATCAA	433
OY	101	GlnAspLeuValThrProAlaAlaAlaAlaProGlnAlaValLysSerSerIleAlaAla	120
DB	434	CAAGACCTGGTTACCCCGCTGACCTGCCCTCAGGCGATGAATCCTATCGCAGCT	493
OY	121	ProSerArgProGluProAsnArgCyseGlySerCyseArgValGlyLeuThrGly	140
DB	494	CCCTTAGACCCGAGCCCAATCGATCGGATCTTCGAGGAACCGTGGATGACAGA	553
OY	141	PheLysCyseArgCyseGlyAsnLeuTyrCyseAlaLeuHisArgTyrSerAspIleThr	160
DB	554	TTTAAGTGTGCTGTGGCAACCTTACTGCGTTTACATCGGTACTCGGACAAACACT	613
OY	161	CysThrTyrAspTyrIleValAlaAlaGlyGlnGluAlaIleAlaLysAlaAsnProLeuVal	180
DB	614	TGCACATATGACTACAAAGCCGAGGGCAGAGGATTTGCAAAAGCTTATCTTGTGC	673
OY	181	ValAlaGluLysValValLysPhe	188
DB	674	GTCGCCGAGAGGTGTCTCAAGTTT	697
RESULT 2			
LOCUS	AX288141	804 bp	DNA linear PAT 21-NOV-2001
DEFINITION	Sequence 12 from Patent WO0177311.		
ACCESSION	AX288141		
VERSION	AX288141.1	GI:17049843	
KEYWORDS	Phycomitrella patens		
SOURCE	Phycomitrella patens		
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta; Bryopsida; Funariidae; Funariales; Funariaceae; Phycomitrella.		
REFERENCE	1 da Costa Silva,O., Bohmert,H.U., van Thielen,N. and Chen,R. Transcription factor stress-related proteins and methods of use in plants		
AUTHORS	Patent: WO 0177311-A 12 18-OCT-2001;		
TITLE	Plants		
JOURNAL	BASF Plant Science GmbH (DE)		
FEATURES	Location/Qualifiers		
source	1..804		
	/organism="Phycomitrella patens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:3218"		
ORIGIN			
Alignment Scores:			
Pred. No.:	6.32e-61	Length:	804
Score:	984.00	Matches:	188

Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0
US-10-716-089-20 (1-188) x AX288141 (1-804)			
OY	1	MetAlaThrGluArgValSerGlnGluThrThrSerGlnAlaProGluG	20
DB	134	ATGGCCACCGAGCGTGTCTCAGAGACGACCTGGCAGGCCCTCGAGG	193
OY	21	CysLysAsnLeuCyseGlyPhePheGlySerGlnAlaThrMetGlyLeuCy	40
DB	194	TGCAAGAACCTTTGCGGCTTCTTCGGCAGCCAACTGAGGGGTGT	253
OY	41	TyrArgGluThrValMetGlnAlaLysMetThrAlaLeuAlaGluGln	60
DB	254	TACCGAGACATGATGACGAGCAAGATGACGGCTTTAGCTGACGACAG	313
OY	61	AlaGlnAlaThrSerAlaThrAlaAlaValAlaGlnProProAlaPro	80
DB	314	GCTCAGGCGACATCTGCCACAGCTGCTGTTTCAGCCCCCTCCTCG	373
OY	81	LysLeuThrCyseGluValGluArgThrMetIleValProHisGlnSer	100
DB	374	AAGCTCACATGCGAGGTGAGAGAACATGATTGTGCCATCAATCTT	433
OY	101	GlnAspLeuValThrProAlaAlaAlaAlaProGlnAlaValLysSer	120
DB	434	CAAGACCTGGTTACCCCGCTGACCTGCCCTCAGGCGATGAATCCTC	493
OY	121	ProSerArgProGluProAsnArgCyseGlySerCyseArgValG	140
DB	494	CCCTTAGACCCGAGCCCAATCGATCGGATCTTCGAGGAACCGTGTG	553
OY	141	PheLysCyseArgCyseGlyAsnLeuTyrCyseAlaLeuHisArgTyrSer	160
DB	554	TTTAAGTGTGCTGTGGCAACCTTACTGCGCTTACATCGGTACTCGG	613
OY	161	CysThrTyrAspTyrIleValAlaAlaGlyGlnGluAlaIleAlaLysAla	180
DB	614	TGCACATATGACTACAAAGCCGAGGGCAGAGGATTTGCAAAAGCTT	673
OY	181	ValAlaGluLysValValLysPhe	188
DB	674	GTCGCCGAGAGGTGTCTCAAGTTT	697
RESULT 3			
LOCUS	AR452942	1531 bp	DNA linear
DEFINITION	Sequence 4 from patent US 6677504.		
ACCESSION	AR452942		
VERSION	AR452942.1	GI:42685089	
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 1531)		
AUTHORS	da Costa e Silva,O., Bohmert,H.U., van Thielen,N. a		
TITLE	Transcription factor stress-related proteins and me		
JOURNAL	Patent: US 6677504-A 13-JAN-2004;		
FEATURES	BASF Plant Science GmbH; Ludwigshafen;		
source	DEX;		
	Location/Qualifiers		
	1..1531		
	/organism="unknown"		
	/mol_type="genomic DNA"		
ORIGIN			
Alignment Scores:			
Pred. No.:	1.03e-58	Length:	1531
Score:	957.00	Matches:	187

Percent Similarity: 99.47%
Best Local Similarity: 99.47%
Query Match: 97.26%
DB: 6
Gaps: 0

Conservative: 0
Matches: 1
Indels: 1
Gaps: 0

US-10-716-089-20 (1-188) x AR452942 (1-1531)

QY 1 MetAlaThrGluArgValSerGlnGluThrThrSerGlnAlaProGluGlyProValMet 20
DB 655 ATGGCCACCGAGGCTGTCTTCAGAGAGACGACCTCCAGGCCCTTAGGGTCCAGTTAAG 714
QY 21 CysLysAsnLeuCyGlyPhePheGlySerGlnAlaThrMetGlyLeuCySerLysGly 40
DB 715 TGCAGAACCTTTGGCGCTTCTTCGCGACGCCAAGCTAACATGGGTTGCTGAAAGTGC 774
QY 41 TTTAGGluThrValMetGlnAlaLysMetThrAlaLeuAlaGluGlnAlaThrGlnAla 60
DB 775 TACCGAGAGACAGTCATGACAGCG-AAAGATGACGGCTTTAGCTGACAGCACTCAGGCT 833
QY 61 AlaGlnAlaThrSerAlaThrAlaAlaAlaValGlnProProAlaProValHISGluThr 80
DB 834 GCTCAGGCGACATGCGCACAGCTGCTGCTTCACGCCCTCTCTGACATGAGACC 893
QY 81 LysLeuThrCyGluGluValGluArgThrMetIleValProHISGlnSerSerSerTyrGln 100
DB 894 AAGCTCACATGCGAGGTTGAGAGAACATGATTTGCGCCGATCAATCTTCAGCTATCA 953
QY 101 GlnAspLeuValThrProAlaAlaAlaAlaProGlnAlaValLysSerSerIleAlaAla 120
DB 954 CAAGACCTGGTTACCCCGCTGACGCTGCCCTCAGGAGTGAAAGTCTTATCCGAGCT 1013
QY 121 ProSerArgProGluProAsnArgCyGlySerCyArgLysArgValGlyLeuThrGly 140
DB 1014 CCTCTAGACCCGAGGCCAATGATCGGATCTTGGAGAGAGCGTTGGATTGACAGCA 1073
QY 141 PheLysCyArgCyGlyAsnLeuTyrCyAlaLeuHisArgTyrSerAspLysHisThr 160
DB 1074 TTTAAGTGTGCTGTGGCAACCTCTTACGCTTTCATCGGTAATCGAGCAAAACACT 1133
QY 161 CysThrTyrAspTyrLysAlaAlaGlyGlnGluAlaIleAlaLysAlaAsnProLeuVal 180
DB 1134 TGCACATATGACTACAAAGCCGACGAGGAGAAAGCATTTGCCAAACTATCTCTTGC 1193
QY 181 ValAlaGluLysValValLysPhe 188
DB 1194 GTGGCCGAGAAAGTTGTCAAGTTT 1217

RESULT 4
AX288133
LOCUS AX288133 1531 bp DNA linear PAT 21-NOV-2001
DEFINITION Sequence 4 from Patent WO0177311.
ACCESSION AX288133
VERSION AX288133.1 GI:17049835
KEYWORDS
SOURCE
ORGANISM Physcomitrella patens
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella.
REFERENCE
AUTHORS da Costa Silva, O., Bohnert, H.J., van Thiel, N., and Chen, R.
TITLE Transcription factor stress-related proteins and methods of use in plants
JOURNAL Patent: WO 0177311-A 4 18-OCT-2001;
FEATURES
source BASF Plant Science GmbH (DE)
Location/Qualifiers
1..1531
/organism="Physcomitrella patens"
/mol_type="unassigned DNA"
/db_xref="taxon:3218"

ORIGIN
Alignment Scores:
Pred. No.: 1.03e-58 Length: 1531

Score: 957.00
Percent Similarity: 99.47%
Best Local Similarity: 99.47%
Query Match: 97.26%
DB: 6
Gaps: 0

Matches: 187
Conservative: 0
Matches: 1
Indels: 1
Gaps: 0

US-10-716-089-20 (1-188) x AX288133 (1-1531)

QY 1 MetAlaThrGluArgValSerGlnGluThrThrSerGlnAlaProGluGly 20
DB 655 ATGGCCACCGAGGCTGTCTTCAGAGAGACGACCTCCAGGCCCTTAGGGTCCAGTTAAG 714
QY 21 CysLysAsnLeuCyGlyPhePheGlySerGlnAlaThrMetGlyLeuCySerLysGly 40
DB 715 TGCAGAACCTTTGGCGCTTCTTCGCGACGCCAAGCTAACATGGGTTGCTGAAAGTGC 774
QY 41 TTTAGGluThrValMetGlnAlaLysMetThrAlaLeuAlaGluGlnAlaThrGlnAla 60
DB 775 TACCGAGAGACAGTCATGACAGCG-AAAGATGACGGCTTTAGCTGACAGCAAC 833
QY 61 AlaGlnAlaThrSerAlaThrAlaAlaAlaValGlnProProAlaProValHISGluThr 80
DB 834 GCTCAGGCGACATGCGCACAGCTGCTGCTTCACGCCCTCTCTGACATGAGACC 893
QY 81 LysLeuThrCyGluGluValGluArgThrMetIleValProHISGlnSerSerSerTyrGln 100
DB 894 AAGCTCACATGCGAGGTTGAGAGAACATGATTTGCGCCGATCAATCTTCAGCTATCA 953
QY 101 GlnAspLeuValThrProAlaAlaAlaAlaProGlnAlaValLysSerSerIleAlaAla 120
DB 954 CAAGACCTGGTTACCCCGCTGACGCTGCCCTCAGGAGTGAAAGTCTTATCCGAGCT 1013
QY 121 ProSerArgProGluProAsnArgCyGlySerCyArgLysArgValGlyLeuThrGly 140
DB 1014 CCTCTAGACCCGAGGCCAATGATCGGATCTTGGAGAGAGCGTTGGATTGACAGCA 1073
QY 141 PheLysCyArgCyGlyAsnLeuTyrCyAlaLeuHisArgTyrSerAspLysHisThr 160
DB 1074 TTTAAGTGTGCTGTGGCAACCTCTTACGCTTTCATCGGTAATCGGTAATCGG 1133
QY 161 CysThrTyrAspTyrLysAlaAlaGlyGlnGluAlaIleAlaLysAlaAsnProLeuVal 180
DB 1134 TGCACATATGACTACAAAGCCGACGAGGAGAAAGCATTTGCCAAACTATCTCTTGC 1193
QY 181 ValAlaGluLysValValLysPhe 188
DB 1194 GTGGCCGAGAAAGTTGTCAAGTTT 1217

RESULT 5
AKI04562
LOCUS AKI04562 999 bp mRNA linear
DEFINITION Oryza sativa (japonica cultivar-group) cDNA clone: Oryza sativa (japonica cultivar-group) cDNA clone: Oryza sativa (japonica cultivar-group) insert sequence.
ACCESSION AKI04562
VERSION AKI04562.1 GI:32989771
KEYWORDS FLI cDNA; oligo capping.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; I
REFERENCE
AUTHORS The Rice Full-Length cDNA Consortium, National Inst
Agrobiological Sciences Rice Full-Length cDNA Proj
Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Do
Kishimoto, N., Yarak, J., Ishikawa, M., Yamada, H., Oo
Kojima, K., Nemiki, T., Ohneda, E., Yahagi, W., Suzuki,
Onseuki, K., Shieniki, T., Foundation of Advancement
Science Genome Sequencing & Analysis Group; Odom,
Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda
Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xi
Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., N
Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., M

Kusunegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN; Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arikawa, T., Fukuda, S., Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kondo, S., Kono, H., Miyazaki, A., Oasato, N., Oe, Y., Saito, R., Sasaki, D., Sato, K., Shiba, K., Shingawa, A., Shiraki, T., Yoshino, M. and Hayashizaki, Y.

TITLE	JOURNAL	PUBMED	REFERENCE	AUTHORS
<p>1869764</p> <p>2 (bases 1 to 1000)</p> <p>Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K., Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Haebizume,W., Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hizemoto,K., Hirooka,T., Hori,F., Horita,I., Iida,J., Iida,Y., Ikeda,R., Imamura,K., Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawagashira,N., Kawai,Y., Kawamata,M., Kikuchi,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M., Kodama,T., Kojima,K., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koye,S., Kurihara,C., Kurotaki,T., Kusumegi,T., Li,C., Lu,M., Maeda,H., Matsubara,K., Matsuyama,T., Miura,J., Miyazaki,A., Mizuno,K., Murakami,K., Murata,M., Nagata,T., Nakamura,M., Nemiki,T., Nariikawa,R., Nishikawa,J., Nishii,K., Nomura,K., Numasaki,R., Ohneda,E., Ohno,M., Ohtsuki,K., Oka,M., Ooka,H., Osato,N., Oca,Y., Ocomi,Y., Ryu,R., Satoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Sato,K., Shibata,K., Shingawa,A., Shiraki,T., Shishiki,T., Sogabe,Y., Sugano,S., Sugiyama,A., Suzuki,K., Suzuki,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toyota,T., Tsunoda,Y., Ueda,M., Waki,K., Xie,Q., Yanagi,W., Yamada,H., Yamamoto,M., Yasunishi,A., Yasaki,J., Yokomizo,S. and Yoshimura,A.</p> <p>Direct Submission</p> <p>Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of Agricultural Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression, 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp, Tel:81-23-838-7007, Fax:81-23-838-7007)</p> <p>This clone is one of the 28K full-length cDNA clones from japonica rice.</p> <p>URL: http://cdna01.dna.affrc.go.jp/cDNA/NIAS_Rice_Full-length_cDNA_Project_Team_kikuchi.S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yasaki,J., Ishikawa,M., Yamada,H., Ooka,H., Horita,I., Kojima,K., Nemiki,T., Ohneda,E., Yanagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T. and Yamamoto,M.</p> <p>FAIS Genome Sequencing & Analysis Group: Ocomi,Y., Iida,Y., Fujimura,T., Ikeda,R., Ishibiki,J., Kawamata,M., Kobayashi,M., Kodama,T., Kurotaki,T., Kusumegi,T., Lu,M., Maeda,H., Miura,J., Mizuno,K., Nariikawa,R., Nishikawa,J., Oka,M., Ryu,R., Sugano,S., Sugiyama,A., Suzuki,Y., Tsunoda,Y., Ueda,M., Xie,Q., Yokomizo,S., Yoshimura,A., Matsubara,K. and Murakami,K.</p> <p>Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hanagaki,T., Hara,A., Haebizume,W., Hayashida,K., Hayatsu,N., Hizemoto,K., Hirooka,T., Hori,F., Horita,I., Iida,J., Iida,Y., Ishikawa,M., Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawagashira,N., Kawai,Y., Kawamata,M., Kikuchi,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M., Kodama,T., Kojima,K., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koye,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishii,K., Nomura,K., Numasaki,R., Ohno,M., Osato,N., Oca,Y., Satoh,K., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Shibata,K., Shingawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toyota,T., Waki,K., Yasunishi,A. and Hayashizaki,Y.</p> <p>location/Qualifiers</p> <p>1..1000</p> <p>/organism="Oryza sativa (japonica cultivar-group)"</p> <p>/mol_type="mRNA"</p> <p>/cultivar="Nipponbare"</p> <p>/db_xref="taxon:39947"</p> <p>/clone="006-301-C11"</p>				

ORIGIN	Alignment Scores:	Pred. No.:	Score:	Percent Similarity:	Best Local Similarity:	Query Match:	DB:
	6.89e-21	Length:	1000	Matches:	82	Conservative: 26	
	419.50	Matches:	82	Conservative: 26		Mismatches: 59	
	60.00%	Conservative: 26		Mismatches: 59		Indels: 13	
	45.56%	Mismatches: 59		Indels: 13		Gaps: 2	
	42.63%	Indels: 13		Gaps: 2			
	15	Gaps: 2					
US-10-716-089-20 (1-188) x AK104175 (1-1000)							
Qy	8	GlnGluThrThrsSerGlnAlaProGlnGluYProValMetCysIysAsnL	27				
Db	160	AAGAGACTGGATGCCAGCAGCCAGAAAGCCGATCTTGTGCATCATATA	219				
Qy	28	PheGlySerGlnAlaThrMetGlyLeuCysSerIysCysThrArgGluT	47				
Db	220	TTTGGCAGTGCCTGCTCCATCAATGACATGTGCTCAAGTCCACAAGAGA	279				
Qy	48	AlAlaMetThrAlaLeuAlaGlnGlnAlaThrGlnAlaAlaGlnAlaT	67				
Db	280	CAGAGACAGGCCAAGCTGGCA-----GCTCTCTATGCACAGCATTC	333				
Qy	68	AlAlaAlaValAlaGlnProProAlaProValHisGluThrIlysLeuThrc	87				
Db	334	GATTCGCCGGAAGAACCATATTATGCTGTCGACAGCCTGAAGTACGTGTG	393				
Qy	88	ArgThrMetIleValProHisGlnSerSerSerTyrGlnGlnAspLeuV	107				
Db	394	GTGAAGACCTGTGT-----G	420				
Qy	108	AlAlaAlaProGlnAlaValIysSerSerIleAlaAlaProSerArgP	127				
Db	421	GAAATTCCTGCTCTCTATGTCAGGGCGCTACCGCTGAACCCACAGGCGG	480				
Qy	128	ArgCysGlySerCysArgIysArgValGlyLeuThrGlyPheIysCysVal	147				
Db	481	CGGTGCTCCACTGTGTCGGAAGAGGGTGTGGCTTATCCGGATTCAATGCC	540				
Qy	148	LeuTyrCysAlaLeuHisArgTyrSerArgAspIysHisThrCysThrIyr	167				
Db	541	TTTACTGCGCAATGCACCGCTATTCGATACCAAGACTCCACAGTTTG	600				
Qy	168	AlaGlyGlnGlnAlaIleAlaIysAlaAspProLeuValValaIaGluT	187				
Db	601	GCTGCTAAGGATGCTATTGTCGAAGCTAATCCGGTGTGAAGCGGAGA	660				
RESULT 7							
AK104605		1009 bp	mrna	linear			
LOCUS	Oryza sativa (japonica cultivar-group)			cDNA clone:O-			
DEFINITION	Insert sequence.						
ACCESSION	AK104605						
VERSION	AK104605.1						
KEYWORDS	FLI-CDNA; oligo capping.						
SOURCE	Oryza sativa (japonica cultivar-group)						
ORGANISM	Oryza sativa (japonica cultivar-group)						
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta						
	Spermatophyta; Magnoliophyta; Liliopsida; Poales; P						
	Pharbitidaceae; Oryzaceae; Oryza.						
REFERENCE							
AUTHORS	1						
	The Rice Full-length cDNA Consortium, National Inst						
	Agrobiological Sciences Rice Full-length cDNA Proj ¹						
	Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Do						
	Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Oo						
	Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,						
	Ohtsuki,K., Shishiki,T., Foundation of Advancement						
	Science Genome Sequencing & Analysis Group:: Ootomo,						
	Iida,Y., Sugano,S., Fujimura,T., Suzuki,T., Tanuda,						
	Kuroseki,T., Kodama,S., Maeda,H., Kobayashi,M., Xi						
	Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., N						
	Ikedo,R., Ishibiki,J., Kawanata,M., Yoshimura,A., N						

TITLE	JOURNAL	PUBLISHED	REFERENCE	AUTHORS
Kusunegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN: Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuoka, S., Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Konno, S., Komuro, H., Miyazaki, A., Osato, N., Oca, Y., Saito, R., Sasaki, D., Sato, K., Shibata, K., Shingawa, A., Shiraki, T., Yoshino, M. and Hayashizaki, Y.	Japanese rice Science 301 (5631), 376-379 (2003)	12869764	(bases 1 to 1009) Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Dol, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hirooka, T., Hori, F., Hotte, I., Iida, J., Iida, Y., Ikeda, R., Immamura, K., Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I., Kawaguchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Koijima, K., Koijima, Y., Konno, S., Komuro, H., Kouda, M., Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M., Nanami, T., Natikawa, R., Niikura, J., Nishi, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H., Osato, N., Oca, Y., Otomo, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K., Sekazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shingawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.	
TITLE	JOURNAL	PUBLISHED	REFERENCE	AUTHORS
Submitted (27-MUG-2002) Shoehi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression, 2-1-2 Kamondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail:shikichi@affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007)	This clone is one of the 28k full-length cDNA clones from japonica rice. URL : http://cdna01.dna.affrc.go.jp/cdna/ NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Dol, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotte, I., Koijima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and Yamamoto, M.	FAIS Genome Sequencing & Analysis Group: Otomo, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Natikawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K.	Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanegawa, S., Katoh, H., Kawai, J., Kishikawa-Hirozane, T., Koijima, Y., Konno, S., Komuro, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakanuma, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N., Oca, Y., Satoh, H., Sakai, C., Sakai, K., Sekazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.	location/Qualifiers 1..1009 /organism="Oryza sativa (japonica cultivar-group)" /mol_type="mRNA" /cui_type="Nipponbare" /db_xref="taxon:39947" /clone="006-309-A12"
ORIGIN	Alignment Scores:	6.96e-21	Length:	1009
	Pred. No.:	419.50	Matches:	82
	Score:	60.00%	Conservative:	26
	Best Local Similarity:	45.56%	Mismatches:	59
	Query Match:	42.63%	Indels:	13
	Gaps:	15	Indels:	2
US-10-716-089-20 (1-188) x AK104605 (1-1009)				
Oy	8 GINGTUTThrSerGlnAlaProGluGlyProValMetCysLysAenL			27
Dd	159 AGGAGACTGCATGCCACAGCGAGAGCCGGATCTTCGATCAATA			218
Oy	28 PhgIySerGlnAlaThrMetGlyLeuCySerLyScyVtyrArgLut			47
Dd	219 TTGTGGCAGGCGGCTCATCATGAACTGTGCTTCAAAGTCCCAAGGA			278
Oy	48 AlAyMetThrAlaLeuAlaGluGlnAlaThrGlnAlaAlaGlnAlat			67
Dd	279 CAGGAGCGAGCCCAAGCTTGCA-----GCTTCCTTCATCACAGATTG			332
Oy	68 AlaAlaAlaValGlnProProAlaProValHisGluThrLysLeuThrc			87
Dd	333 GATTCGCGGAGAAGAACCAATTATTTCTGTCGTCGATCGATGAGCTGTG			392
Oy	88 ArgThrMetIleValProHisGlnSerSerSerTyrgInglNaPleuv			107
Dd	393 GTGAGAGACCGTGTG-----GCTTCCTTCATCACAGATTG			419
Oy	108 AlaAlaAlaProGlnAlaValLysSerIleAlaAlaProSerArgP			127
Dd	420 GAGATTGCTGCTCCTAGTAGAGGGGCTGTAACCCCAAGGGAGGG			479
Oy	128 ATGCGGGLySerCyArgLysArgValGlyLeuThrGlyPheLysCyga			147
Dd	480 CGGTGCTCCACTGTTCGAGAGAGGCTGTGCTTACCAGGATTCACCTGC			539
Oy	148 LeuTYRCyaaLLeuHiARGYTSerAspLysIsthrCysThrTYZA			167
Dd	540 TTGTACTGGCAATGACACCGCATTTCCATTAAGCATGATCCTCAGTTTG			599
Oy	168 AlaGlyGlnGlnAlaIleAlaLysAlaAnProLeuValAlaAlaGluL			187
Dd	600 GCTGCTAGGATGCTATTGCGCAAGGCTAATCGGTGCGAAGCGGAGA			659
RESULT 8	AK060008	1021 bp	mRNA	linear
LOCUS	Oryza sativa (japonica cultivar-group) cDNA clone:01			
DEFINITION	Insert sequence.			
ACCESSION	AK060008			2003
VERSION	AK060008.1 GI:32970026			Full
SOURCE	FLI_CDNA: oligo-capping			
ORGANISM	Oryza sativa (japonica cultivar-group)			
REFERENCE	Oryza sativa (japonica cultivar-group) Embryophyta			
AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Magnoliopsida; Poales; P			
	Bhrnatoidae; Oryzeae; Oryza.			
	The Rice Full-length cDNA Consortium, National Inst			
	Agrobiological Sciences Rice Full-length cDNA Proje			
	Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Do			
	Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Oo			
	Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki,			
	Ohtsuki, K., Shishiki, T., Foundation of Advancem			
	Science Genome Sequencing & Analysis Group: Otomo,			
	Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda,			
	Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xi			
	Natikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., N			
	Ikeda, R., Ishibiki, J., Kawamata, M., Yamamoto, A., M			

Kusunomegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., Riken, K., Kawa, J., Carninci, P., Adachi, J., Aizawa, K., Arikawa, T., Fukuda, S., Hara, A., Hashidume, W., Hayashi, N., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kondo, S., Komuro, H., Miyazaki, A., Oosato, N., Oda, Y., Saito, R., Sasaki, D., Sato, K., Shibata, K., Shingawa, A., Shiraki, T., Yoshino, M., and Hayashizaki, Y.

JOURNAL
Science 301 (5631), 376-379 (2003)

PUBMED 12869764

2 (BASES 1 TO 1021)
Adachi, T. Aizawa, K.

Fujimura, T., Fukuda, S., Hanagaki, T., Harata, A., Hashizume, W., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hirooka, T., Horii, P., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K., Imotani, K., Ishibiki, J., Ienli, Y., Ishikawa, M., Itoh, M., Kageawa, I., Kanegawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M., Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Komoto, H., Kouda, M., Koya, S., Kurihara, C., Kurossaki, T., Kusumegi, T., Ii, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M., Namiki, T., Narikawa, R., Nikkura, Y., Nishi, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H., Osato, N., Ota, Y., Otomo, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takehashi, F., Takaku-Akhiba, S., Tanaka, T., Tomaru, A., Toyota, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.

TITLE Direct Submission
JOURNAL Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of
Agrobiological Sciences, Department of Molecular Genetics, Head of
Laboratory of Gene Expression, 2-1-2 Kannondai, Tsukuba, Ibaraki
305-8603, Japan (E-mail:skikuchi@nias.affrc.go.jp,
Tel:81-29-838-7007, Fax:81-29-838-7007)
COMMENT This clone is one of the 28K full-length cDNA clones from japonica
rice.

FEATURES	Location/Qualifiers
SOURCE	1. .1021

```

/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="caxon:33947"
/clone="06-302-C08"

```

ORIGIN

Alignment Scores:	
Pred. No.:	7.05e-21
Score:	419.50
Percent Similarity:	60.00%
Best Local Similarity:	45.56%
Query Match:	42.63%
DB:	15
Gaps:	2
Length:	1021
Matches:	82
Conservative:	26
Mismatches:	59
Indels:	13

US-10-716-089-20 (1-188) x AK060008 (1-1021)

[illegible]

Science Genome Sequencing & Analysis Group: Ootomo, Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Kurotaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xi Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., N Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., M

TITLE	JOURNAL PUBMED REFERENCE AUTHORS
Kuunemegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN, Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Oeato, N., Oca, Y., Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Yoshino, M., and Hayashizaki, Y. Collection, mapping, and annotation of over 28,000 cDNA clones from Japanese rice Science 301 (5631), 376-379 (2003)	12869764 2 (bases 1 to 1209)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K., Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M., Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Kuroaki, T., Kusumegi, T., Li, C., Lu, M., Maeda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Miuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M., Naniki, T., Narikawa, R., Niihara, J., Nishi, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H., Oeato, N., Oca, Y., Oeato, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Tsumoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.	Submitted (05-DEC-2001) Shosht Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression, 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007) This clone is one of the 28k full-length cDNA clones from japonica rice. URL : http://cdna01.dna.affrc.go.jp/cDNA/ N1S Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and Yamamoto, M.
FAIS Genome Sequencing & Analysis Group: Oeato, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kuroaki, T., Kuunemegi, T., Lu, M., Maeda, H., Miura, J., Miuno, K., Narikawa, R., Niihara, J., Nomura, K., Nomura, S., Sugiyama, A., Suzuki, Y., Tsumoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K. Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, T., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Oeato, N., Oca, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.	Location/Qualifiers 1..1209 /organism="Oryza sativa (japonica cultivar-group)" /mol_type="mRNA" /cuiovar="Nippondare" /db_xref="taxon:39947" /clone="J023026K20"

ORIGIN	US-10-716-089-20 (1-188) x AK069720 (1-1209)
Alignment Scores: Pred. No.: Score: Percent Similarity: Best Local Similarity: Query Match:	8,46e-21 419.50 60.00% 45.56% 42.63%
DB:	15 Gaps: 2
Oy	8 GlnGlnThrThrserrGlnAlaProGluGlyProValMetCyblyAsnL
Db	157 AAGAGACGTGGATCCAGCGACCGGAGGCCGATCTTCATCAATAT
Oy	28 PheGlySerGlnAlaThrMetGlyLeuCybSerIysCybTyraGluTl
Db	217 TTTCGACATCGCGCTACCATGATGCTCTCAAGTCCCAAGAGAG
Oy	48 AlaIysMetThrAlaLeuAlaGluGlnAlaThrGlnAlaGlnAlaT
Db	277 CAGGAGCAGGCCAAGCTGGCA-----GCTCTCTTATCCACAGATTG
Oy	68 AlaAlaAlaValGlnProProAlaProValHisGluThrIlyLeuThrC
Db	331 GATTCCGGGAAGAACCAATTATTCGTGCTGTCACCGCTGAAGTGTGG
Oy	88 ArgThrMetIleValProHisGlnSerSerSerIyGlnGlnApleuV
Db	391 GTGAAGACGCTTGT-----G
Oy	108 AlaAlaAlaProGlnAlaValIlySerSerIleAlaAlaProSerArgP
Db	418 GAGATTGCTGGCCCTTAGTGAAGGGGTCAAGTGAACCCAGGAGGG
Oy	128 ArgCyblySerCybArgIlyArgValGlyLeuThrGlyPheIlyCybA
Db	478 CGGATCTCCACTTGTCCGAAGAGGTTGCTTACCGGATTCACATGTC
Oy	148 LeuTyrcyAlaLeuHisIleArgTySerAspIlyHisIthrCybThrTyra
Db	538 TTGTACTGCGGCAAGCAACCGCTATTCGATACATGACATGCCAGTTTG
Oy	168 AlaGlyGlnGlnAlaIleAlaIlyAlaAsnProLeuValAlaGluL
Db	598 GCTGCTAGGATGCTATTGCTCAAGGCTTATCCGTTGGTGAAGCGGAGA
RESULT 10	
AP008212_240	
WPCOMMENT	
Sequence split into 308 fragments LOCUS AP008212 Accession A1	
Fragment Name	Begin End
AP008212_000	1 110000
AP008212_001	100001 210000
AP008212_002	200001 310000
AP008212_003	300001 410000
AP008212_004	400001 510000
AP008212_005	500001 610000
AP008212_006	600001 710000
AP008212_007	700001 810000
AP008212_008	800001 910000
AP008212_009	900001 1010000
AP008212_010	1000001 1110000
AP008212_011	1100001 1210000
AP008212_012	1200001 1310000
AP008212_013	1300001 1410000
AP008212_014	1400001 1510000
AP008212_015	1500001 1610000
AP008212_016	1600001 1710000
AP008212_017	1700001 1810000
AP008212_018	1800001 1910000
AP008212_019	1900001 2010000

AP008212_020 2000001 2110000
AP008212_021 2100001 2210000
AP008212_022 2200001 2310000
AP008212_023 2300001 2410000
AP008212_024 2400001 2510000
AP008212_025 2500001 2610000
AP008212_026 2600001 2710000
AP008212_027 2700001 2810000
AP008212_028 2800001 2910000
AP008212_029 2900001 3010000
AP008212_030 3000001 3110000
AP008212_031 3100001 3210000
AP008212_032 3200001 3310000
AP008212_033 3300001 3410000
AP008212_034 3400001 3510000
AP008212_035 3500001 3610000
AP008212_036 3600001 3710000
AP008212_037 3700001 3810000
AP008212_038 3800001 3910000
AP008212_039 3900001 4010000
AP008212_040 4000001 4110000
AP008212_041 4100001 4210000
AP008212_042 4200001 4310000
AP008212_043 4300001 4410000
AP008212_044 4400001 4510000
AP008212_045 4500001 4610000
AP008212_046 4600001 4710000
AP008212_047 4700001 4810000
AP008212_048 4800001 4910000
AP008212_049 4900001 5010000
AP008212_050 5000001 5110000
AP008212_051 5100001 5210000
AP008212_052 5200001 5310000
AP008212_053 5300001 5410000
AP008212_054 5400001 5510000
AP008212_055 5500001 5610000
AP008212_056 5600001 5710000
AP008212_057 5700001 5810000
AP008212_058 5800001 5910000
AP008212_059 5900001 6010000
AP008212_060 6000001 6110000
AP008212_061 6100001 6210000
AP008212_062 6200001 6310000
AP008212_063 6300001 6410000
AP008212_064 6400001 6510000
AP008212_065 6500001 6610000
AP008212_066 6600001 6710000
AP008212_067 6700001 6810000
AP008212_068 6800001 6910000
AP008212_069 6900001 7010000
AP008212_070 7000001 7110000
AP008212_071 7100001 7210000
AP008212_072 7200001 7310000
AP008212_073 7300001 7410000
AP008212_074 7400001 7510000
AP008212_075 7500001 7610000
AP008212_076 7600001 7710000
AP008212_077 7700001 7810000
AP008212_078 7800001 7910000
AP008212_079 7900001 8010000
AP008212_080 8000001 8110000
AP008212_081 8100001 8210000
AP008212_082 8200001 8310000
AP008212_083 8300001 8410000
AP008212_084 8400001 8510000
AP008212_085 8500001 8610000
AP008212_086 8600001 8710000
AP008212_087 8700001 8810000
AP008212_088 8800001 8910000
AP008212_089 8900001 9010000
AP008212_090 9000001 9110000
AP008212_091 9100001 9210000
AP008212_092 9200001 9310000

AP008212_093 9300001 9410000
AP008212_094 9400001 9510000
AP008212_095 9500001 9610000
AP008212_096 9600001 9710000
AP008212_097 9700001 9810000
AP008212_098 9800001 9910000
AP008212_099 9900001 10010000
AP008212_100 1000001 10110000
AP008212_101 1010001 10210000
AP008212_102 1020001 10310000
AP008212_103 1030001 10410000
AP008212_104 1040001 10510000
AP008212_105 1050001 10610000
AP008212_106 1060001 10710000
AP008212_107 1070001 10810000
AP008212_108 1080001 10910000
AP008212_109 1090001 11010000
AP008212_110 1100001 11110000
AP008212_111 1110001 11210000
AP008212_112 1120001 11310000
AP008212_113 1130001 11410000
AP008212_114 1140001 11510000
AP008212_115 1150001 11610000
AP008212_116 1160001 11710000
AP008212_117 1170001 11810000
AP008212_118 1180001 11910000
AP008212_119 1190001 12010000
AP008212_120 1200001 12110000
AP008212_121 1210001 12210000
AP008212_122 1220001 12310000
AP008212_123 1230001 12410000
AP008212_124 1240001 12510000
AP008212_125 1250001 12610000
AP008212_126 1260001 12710000
AP008212_127 1270001 12810000
AP008212_128 1280001 12910000
AP008212_129 1290001 13010000
AP008212_130 1300001 13110000
AP008212_131 1310001 13210000
AP008212_132 1320001 13310000
AP008212_133 1330001 13410000
AP008212_134 1340001 13510000
AP008212_135 1350001 13610000
AP008212_136 1360001 13710000
AP008212_137 1370001 13810000
AP008212_138 1380001 13910000
AP008212_139 1390001 14010000
AP008212_140 1400001 14110000
AP008212_141 1410001 14210000
AP008212_142 1420001 14310000
AP008212_143 1430001 14410000
AP008212_144 1440001 14510000
AP008212_145 1450001 14610000
AP008212_146 1460001 14710000
AP008212_147 1470001 14810000
AP008212_148 1480001 14910000
AP008212_149 1490001 15010000
AP008212_150 1500001 15110000
AP008212_151 1510001 15210000
AP008212_152 1520001 15310000
AP008212_153 1530001 15410000
AP008212_154 1540001 15510000
AP008212_155 1550001 15610000
AP008212_156 1560001 15710000
AP008212_157 1570001 15810000
AP008212_158 1580001 15910000
AP008212_159 1590001 16010000
AP008212_160 1600001 16110000
AP008212_161 1610001 16210000
AP008212_162 1620001 16310000
AP008212_163 1630001 16410000
AP008212_164 1640001 16510000
AP008212_165 1650001 16610000

AP008212_166 16600001 16710000
AP008212_167 16700001 16810000
AP008212_168 16800001 16910000
AP008212_169 16900001 17010000
AP008212_170 17000001 17110000
AP008212_171 17100001 17210000
AP008212_172 17200001 17310000
AP008212_173 17300001 17410000
AP008212_174 17400001 17510000
AP008212_175 17500001 17610000
AP008212_176 17600001 17710000
AP008212_177 17700001 17810000
AP008212_178 17800001 17910000
AP008212_179 17900001 18010000
AP008212_180 18000001 18110000
AP008212_181 18100001 18210000
AP008212_182 18200001 18310000
AP008212_183 18300001 18410000
AP008212_184 18400001 18510000
AP008212_185 18500001 18610000
AP008212_186 18600001 18710000
AP008212_187 18700001 18810000
AP008212_188 18800001 18910000
AP008212_189 18900001 19010000
AP008212_190 19000001 19110000
AP008212_191 19100001 19210000
AP008212_192 19200001 19310000
AP008212_193 19300001 19410000
AP008212_194 19400001 19510000
AP008212_195 19500001 19610000
AP008212_196 19600001 19710000

Alignment Scores:

Pred. No.: 1,116-18 Length: 110000
Score: 419.50 Matches: 82
Percent Similarity: 60.00% Conservative: 26
Best Local Similarity: 45.56% Mismatches: 59
Query Match: 42.63% Indels: 13
DB: 15 Gaps: 2

US-10-716-089-20 (1-188) x AP008212_240 (1-110000)

QY 8 GInGIuThrThSerGInAlaIeValProValMetCyAlaAsnLeuCySgIyPhe 27
DB 109415 AAGGAGACTGCATGCGACGACGACGACCCGATCCTTGATCATTAATCGCGCTTC 109474
QY 28 PheGIySerGInAlaIeThrMetGlyLeuCySerIySvYrYrAsgIuThrValMetGln 47
DB 109475 TTTGGCAGTGGCGGTACCATGAAATGTCCTCAAGTGCACAGAGATGATTAATGAAG 109534
QY 48 AlaIySerMetThrAlaIeValaGluGlnAlaIeThrGlnAlaIaGlnAlaIeThrSerIaThr 67
DB 109535 CAGGAGCAGGCCAGCTGCA-----GCCTCCTCATCGACAGCATTTGCTCAATGTGCT 109588
QY 68 AlaAlaAlaValaGlnProProAlaProValaHisGluThrIyLeuThrCySgIyValaGlu 87
DB 109589 GATTCCGGGAGAGAACCAATTATTGCTGTCACGCTGAAGTACGTGCTCAATCGAG 109648
QY 88 ArgThrMetIleValProHisGlnSerSerSerTyrgInGlnAspLeuValThrProAla 107
DB 109649 GTGAAGACGCTGTGT-----GCCGAGCCTGCT 109675
QY 108 AlaAlaAlaProGlnAlaValIySerSerIeAlaIaProSerArypProGluProAsn 127
DB 109676 GAGATTGCTGCTCCTAGTGAAGGGGTCAACGTGAACCCAGAGGGAGAGACCAAT 109735
QY 128 ArgCySgIySerCyArygIyAryValaGlyLeuThrGlyPheIyCyAryCySgIyAsn 147
DB 109736 CGGTCTCTCCACTTGTTCGAAAGGGGTGTCTTACCGGATTCACCTGCGATGTGCGAAC 109795
QY 148 LeuTyCyValaLeuHisIeArgTySerArypIyShIeThrCySvThrTyTrAspTyTrIyAla 167
DB 109796 TTGTACTGCGCAATGCACCGCTATTTCGATTAAGCATGACTGCAGTTTGAATATCGAGAC 109855

QY 168 AlaGIyGlnAlaIeAlaIeAlaIyAsnProLeuValValaIaGluIy 187
DB 109856 GCTCTAGGATGCTATTGCCAAGCTTAATCCGGTGTGAAGCGGAG 10991

RESULT 11

AP008212_241

MPCOMMENT

Sequence split into 308 fragments. LOCUS AP008212 Accession A1

Fragment Name	Begin	End
AP008212_000	1	110000
AP008212_001	100001	210000
AP008212_002	200001	310000
AP008212_003	300001	410000
AP008212_004	400001	510000
AP008212_005	500001	610000
AP008212_006	600001	710000
AP008212_007	700001	810000
AP008212_008	800001	910000
AP008212_009	900001	1010000
AP008212_010	1000001	1110000
AP008212_011	1100001	1210000
AP008212_012	1200001	1310000
AP008212_013	1300001	1410000
AP008212_014	1400001	1510000
AP008212_015	1500001	1610000
AP008212_016	1600001	1710000
AP008212_017	1700001	1810000
AP008212_018	1800001	1910000
AP008212_019	1900001	2010000
AP008212_020	2000001	2110000
AP008212_021	2100001	2210000
AP008212_022	2200001	2310000
AP008212_023	2300001	2410000
AP008212_024	2400001	2510000
AP008212_025	2500001	2610000
AP008212_026	2600001	2710000
AP008212_027	2700001	2810000
AP008212_028	2800001	2910000
AP008212_029	2900001	3010000
AP008212_030	3000001	3110000
AP008212_031	3100001	3210000
AP008212_032	3200001	3310000
AP008212_033	3300001	3410000
AP008212_034	3400001	3510000
AP008212_035	3500001	3610000
AP008212_036	3600001	3710000
AP008212_037	3700001	3810000
AP008212_038	3800001	3910000
AP008212_039	3900001	4010000
AP008212_040	4000001	4110000
AP008212_041	4100001	4210000
AP008212_042	4200001	4310000
AP008212_043	4300001	4410000
AP008212_044	4400001	4510000
AP008212_045	4500001	4610000
AP008212_046	4600001	4710000
AP008212_047	4700001	4810000
AP008212_048	4800001	4910000
AP008212_049	4900001	5010000
AP008212_050	5000001	5110000
AP008212_051	5100001	5210000
AP008212_052	5200001	5310000
AP008212_053	5300001	5410000
AP008212_054	5400001	5510000
AP008212_055	5500001	5610000
AP008212_056	5600001	5710000
AP008212_057	5700001	5810000
AP008212_058	5800001	5910000
AP008212_059	5900001	6010000
AP008212_060	6000001	6110000
AP008212_061	6100001	6210000
AP008212_062	6200001	6310000

AP008212_063 6300001 6410000
AP008212_064 6400001 6510000
AP008212_065 6500001 6610000
AP008212_066 6600001 6710000
AP008212_067 6700001 6810000
AP008212_068 6800001 6910000
AP008212_069 6900001 7010000
AP008212_070 7000001 7110000
AP008212_071 7100001 7210000
AP008212_072 7200001 7310000
AP008212_073 7300001 7410000
AP008212_074 7400001 7510000
AP008212_075 7500001 7610000
AP008212_076 7600001 7710000
AP008212_077 7700001 7810000
AP008212_078 7800001 7910000
AP008212_079 7900001 8010000
AP008212_080 8000001 8110000
AP008212_081 8100001 8210000
AP008212_082 8200001 8310000
AP008212_083 8300001 8410000
AP008212_084 8400001 8510000
AP008212_085 8500001 8610000
AP008212_086 8600001 8710000
AP008212_087 8700001 8810000
AP008212_088 8800001 8910000
AP008212_089 8900001 9010000
AP008212_090 9000001 9110000
AP008212_091 9100001 9210000
AP008212_092 9200001 9310000
AP008212_093 9300001 9410000
AP008212_094 9400001 9510000
AP008212_095 9500001 9610000
AP008212_096 9600001 9710000
AP008212_097 9700001 9810000
AP008212_098 9800001 9910000
AP008212_099 9900001 10010000
AP008212_100 10000001 10110000
AP008212_101 10100001 10210000
AP008212_102 10200001 10310000
AP008212_103 10300001 10410000
AP008212_104 10400001 10510000
AP008212_105 10500001 10610000
AP008212_106 10600001 10710000
AP008212_107 10700001 10810000
AP008212_108 10800001 10910000
AP008212_109 10900001 11010000
AP008212_110 11000001 11110000
AP008212_111 11100001 11210000
AP008212_112 11200001 11310000
AP008212_113 11300001 11410000
AP008212_114 11400001 11510000
AP008212_115 11500001 11610000
AP008212_116 11600001 11710000
AP008212_117 11700001 11810000
AP008212_118 11800001 11910000
AP008212_119 11900001 12010000
AP008212_120 12000001 12110000
AP008212_121 12100001 12210000
AP008212_122 12200001 12310000
AP008212_123 12300001 12410000
AP008212_124 12400001 12510000
AP008212_125 12500001 12610000
AP008212_126 12600001 12710000
AP008212_127 12700001 12810000
AP008212_128 12800001 12910000
AP008212_129 12900001 13010000
AP008212_130 13000001 13110000
AP008212_131 13100001 13210000
AP008212_132 13200001 13310000
AP008212_133 13300001 13410000
AP008212_134 13400001 13510000
AP008212_135 13500001 13610000

AP008212_136 13600001 13710000
AP008212_137 13700001 13810000
AP008212_138 13800001 13910000
AP008212_139 13900001 14010000
AP008212_140 14000001 14110000
AP008212_141 14100001 14210000
AP008212_142 14200001 14310000
AP008212_143 14300001 14410000
AP008212_144 14400001 14510000
AP008212_145 14500001 14610000
AP008212_146 14600001 14710000
AP008212_147 14700001 14810000
AP008212_148 14800001 14910000
AP008212_149 14900001 15010000
AP008212_150 15000001 15110000
AP008212_151 15100001 15210000
AP008212_152 15200001 15310000
AP008212_153 15300001 15410000
AP008212_154 15400001 15510000
AP008212_155 15500001 15610000
AP008212_156 15600001 15710000
AP008212_157 15700001 15810000
AP008212_158 15800001 15910000
AP008212_159 15900001 16010000
AP008212_160 16000001 16110000
AP008212_161 16100001 16210000
AP008212_162 16200001 16310000
AP008212_163 16300001 16410000
AP008212_164 16400001 16510000
AP008212_165 16500001 16610000
AP008212_166 16600001 16710000
AP008212_167 16700001 16810000
AP008212_168 16800001 16910000
AP008212_169 16900001 17010000
AP008212_170 17000001 17110000
AP008212_171 17100001 17210000
AP008212_172 17200001 17310000
AP008212_173 17300001 17410000
AP008212_174 17400001 17510000
AP008212_175 17500001 17610000
AP008212_176 17600001 17710000
AP008212_177 17700001 17810000
AP008212_178 17800001 17910000
AP008212_179 17900001 18010000
AP008212_180 18000001 18110000
AP008212_181 18100001 18210000
AP008212_182 18200001 18310000
AP008212_183 18300001 18410000
AP008212_184 18400001 18510000
AP008212_185 18500001 18610000
AP008212_186 18600001 18710000
AP008212_187 18700001 18810000
AP008212_188 18800001 18910000
AP008212_189 18900001 19010000
AP008212_190 19000001 19110000
AP008212_191 19100001 19210000
AP008212_192 19200001 19310000
AP008212_193 19300001 19410000
AP008212_194 19400001 19510000
AP008212_195 19500001 19610000
AP008212_196 19600001 19710000

Alignment Scores:
Pred. No.: 1,11e-18 Length: 110000
Score: 419.50 Matches: 82
Percent Similarity: 60.00% Conservative: 26
Best Local Similarity: 45.56% Mismatches: 59
Query Match: 42.63% Indels: 13
DB: 15 Gaps: 2

US-10-716-089-20 (1-188) x AP008212_241 (1-110000)

Cy 8 GINGUthRthRserGlnAlaPProGluGlyProValMetCysIysAsnL

```

Db      9415 AAGGAGATGATGATGCACAGCAGCAAGCCCGATCCTTGATCAATAACTGGCGCTTC 9474
      28 PheglYSerGlnAlaIthMeTcylLeuCYserlYSyCYeTYrYArgGluThValMeGln 47
Db      9475 TTTGGCAATGGCGGGCTACATGAACTGTGCTCAAGTCCCAAGAGATATATATGAG 9534
      48 AlalYSMeTThrAlaLeuAlaGluInAlaIthrGlnAlaIthGlnAlaIthSerAlaThr 67
Db      9535 CAGGAGCAGCCAGCAGCTGCA-----GCTTCCTTATCGACAGCATTTGCAATGCTGCT 9588
      68 AlAlaAlaValAlnProProAlaProValHisGluThrlyLeuThrCYeGluValGlu 87
Db      9589 GATTCCGGAGAGAACCAATTATTGCTGTGTCACCTGAAGTAGACTGTGCTCAAGTCCAG 9648
      88 ArgThrMeTlleValProHisGlnSerSerSerSerTYrGlnGlnApleuValIthProAla 107
Db      9649 GTGAAGACGCTTGT-----GCGCAGCCTTCT 9675
      108 AlAlaAlaProGlnAlaVallySerSerSerIleAlaIaProSerArgProGluProAen 127
      9676 GAGATTGCTGCTCTAGAGAGGGGTCAAGGTAACCCCAAGGGAGGGAGAGACCAAT 9735
      128 ArgCYeGlySerCYeArglyeArglyeAlglyLeuthrGlyPheLysCYeArgCYeGlyAen 147
Db      9736 CGGTGCTCCACTGTGTCGGAAGAGGGTGTGCTTACCGGATTCACCTGCCGATGTGGCAAC 9795
      148 LeuTYrCYeAlaLeuHisArgTYSerAplelyeHisThrCYeThrTYrAePTyrlYsAla 167
Db      9796 TTGTACTCGGACATGCACCGCTATTCCGATTAAGCATGACTGCGACTTGTACTATCGGACC 9855
      168 AlaGlyGlnGluAlaIleAlaIalysAlaAenProLeuValAlaIaGlyLeuValAllys 187
      9856 GCTGCTAGAGATGCTATTGCTCCAGGCTTAATCCGGTGTGAAGCGGAGAACTTACAAA 9915

RESULT 12
LOCUS   AP003626 158321 bp DNA linear PLN 10-AUG-2004
DEFINITION Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 6,
PAC clone:P0429G06.
ACCESSION AP003626 GI:51090899
SOURCE    Oryza sativa (japonica cultivar-group)
           Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
           Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
           Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1
AUTHORS  Sasaki,T., Matsumoto,T. and Yamamoto,K.
TITLE     Oryza sativa nipponbare (GA3) genomic DNA, chromosome 6, PAC
           clone:P0429G06
JOURNAL   Published Only in Database (2001)
AUTHORS  Sasaki,T., Matsumoto,T. and Yamamoto,K.
TITLE     Direct Submission
JOURNAL   Submitted (17-MAY-2001) Takuji Sasaki, National Institute of
           Agrobiological Sciences, Rice Genome Research Program, Kannondai
           2-1-2, Tsukuba, Ibaraki 305-8602, Japan
           (E-mail:tsasaki@nias.affrc.go.jp, URL:http://ryg.dna.affrc.go.jp/,
           Tel:81-298-38-7441, Fax:81-298-38-7468)
           On Aug 9, 2004 this sequence version replaced gi:47971619.
           Genes were predicted from the integrated results of the following:
           GENSCAN (http://CCR-081.mt.edu/GENSCAN.html), FGENESH
           (http://www.softberry.com/), GeneMark.hmm
           (http://opal.biology.gatech.edu/GeneMark/), GlimmerM
           (http://www.tigr.org/tcb/glimmer/glmr_form.html), RiceHMM
           (http://xip.dna.affrc.go.jp/RiceHMM/), SplicePredictor
           (http://bioinformatics.jaist.ac.jp/cgi-bin/sp.cgi), sim4
           (http://jblab.cse.psu.edu/html/docs/sim4.html), gap2
           (http://www.tigr.org/software/glimmer/), BLASTN and BLASTX. The
           genomic sequence was searched against NCBI Nonredundant Protein
           database, nr (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA

```

FEATURES

source

sequence database at RGP or DBJ. Protein homologue regions were searched against NCBI Nonredundant Pro with BLASTP. ESTs represent the identified cDNA seq BLASTN with the corresponding DBJ accession no. An full-length cDNAs representing the identified cDNA seq BLASTN with the corresponding DBJ accession no. A gene with identity or significant homology to a F classified based on the protein name to indicate th such as same name, 'putative-' and '-like protein'. Significant homology to any protein but with full-1 EST homology (covering almost the entire length of sequence) is classified as an 'unknown' protein. A by two or more gene prediction programs is classifi 'hypothetical' protein according to IREGS standard. predicted by a single gene prediction program is al; a probable 'hypothetical' protein and is included at miscellaneous feature of the sequence. The orientation of the sequence is from SP6 to T7 o This sequence of P0429G06 clone has an overlap with AP003616) clone at 5' end and with P0417G12 (DBJ; / end. Detailed information on overlap and assembly ; with annotation of this entry is available at http://ryg.dna.affrc.go.jp/GenomeSeq.html.

Location/Qualifiers

1. .158321

/organism="Oryza sativa (japonica cultivar

/mol_type="Genomic DNA"

/cultivar="Nipponbare"

/db_xref="taxon:39947"

/chromosome="6"

/clone="P0429G06"

1460. .3909

/gene="P0429G06.1"

join(<1460. .1520,1585. .1912,3783. .>3909

/gene="P0429G06.1"

/note="start and end point are not identifi

join(1460. .1520,1585. .1912,3783. .3909)

/gene="P0429G06.1"

/note="Predicted by GENSCAN etc."

/codon_start=1

/product="hypothetical protein"

/protein_id="BAD35505.1"

/db_xref="GI:51090900"

/translation="MNACCLCTLDGTGVKSRMPGDRSRHSDA

HAGVDPMRMHRAILRASCYAGSVGVAASARMTSOPVA

PPPLMAGSRARILPCRAKYPDFGICGRSPQETVTLFVN

QQKEKINSKD"

5867. .8209

/gene="P0429G06.2"

join(5867. .6041,6152. .6487,6924. .7466,7

/gene="P0429G06.2"

/note="supported by full-length cDNA(s) : A

join(5898. .6041,6152. .6487,6924. .7466,7

/gene="P0429G06.2"

/note="contains EST(s) : C73691(E20165)

contains full-length cDNA(s) : AK119501"

/codon_start=1

/product="putative polygalacturonase"

/protein_id="BAD35506.1"

/db_xref="GI:51090901"

/translation="MPPLDITWSSLLITTSMMWLQATCTARQI

DNSKAFERANAKAAPRAAVVPAAGGCGGCGGYLHV

VVRAPAGLDAPRGHEWINFAGIDGLVYTGCTGDEGASMI

SIKIGSVNRTTGTSTLDSKFEHTVTVGSDVSVHSIRAI

GVRTTDAVATGDCVSVGSRADVTYISGVSCEGHEGISVS

NCTTAGVANGRIKTRWGGRSSAAALAAASGLVEDIVM

LSCHROSERASVVRISDVFRNTRIGVATQVAVKLCSAAS

RGVATVSRCAVAGVAGTVLPPCI"

complement(join(9005. .9081,9543. .9609,97

9853. .9984))

/gene="P0429G06.3"

complement(join(9005. .9081,9543. .9609,97

9853. .9984))

/misc_feature

misc_feature

ing
e
ID.
level
hout
ed
d as
one.
DBJ:
3'
ther

RPOL
SART
IAIG

DDQT
RVAG
PPPS
OGST
LRVS
YCPY
RYVR

SOURCE	Oryza sativa (japonica cultivar-group)
ORGANISM	Oryza sativa (japonica cultivar-group)
REFERENCE	Eukaryota: Viridiplantae: Streptophyta; Tracheophyta; Spermatophyta Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.
AUTHORS	1 Sasaki,T., Matsumoto,T. and Yamamoto,K.
TITLE	Oryza sativa japonica (GAJ) genomic DNA, chromosome 6, PAC clone: P0417G12
JOURNAL	Published Only in Database (2001)
REFERENCE	2 (bases 1 to 163568)
AUTHORS	Sasaki,T., Matsumoto,T. and Yamamoto,K.
JOURNAL	Direct Submission
COMMENT	Submitted (30-MAY-2001) Takuji Sasaki, National Institute of Agrobiological Sciences, Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan (E-mail: tsasaki@nias.affrc.go.jp, URL: http://xgp.dna.affrc.go.jp/, Tel: 81-298-38-7441, Fax: 81-298-38-7468) On Aug 9, 2004 this sequence version replaced gi:48525320. Genes were predicted from the integrated results of the following: GENSCAN (http://CCR-081.mt.edu/GENSCAN.html), FGENESH (http://www.softberry.com/), GeneMark.hmm (http://opal.biology.gatech.edu/GeneMark/), GlimmerM (http://www.tigr.org/cdb/glimmerm/glmr_form.html), RiceHMM (http://xgp.dna.affrc.go.jp/RiceHMM/), SplinePredictor (http://bioinformatics.jastate.edu/cgi-bin/sp.cgi), sim4 (http://globin.cse.psu.edu/html/docs/sim4.html), gap2 (http://www.tigr.org/software/glimmerm/), BLASTN and BLASTX. The genomic sequence was searched against NCBI Nonredundant Protein database, nr (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA sequence database at RGP or DDBJ. Protein homologues of the coding regions were searched against NCBI Nonredundant Protein database with BLASTP. ESTs represent the identified cDNA sequences using BLASTN with the corresponding DDBJ accession no. and RGP clone ID. Full-length cDNAs represent the identified cDNA sequences using BLASTN with the corresponding DDBJ accession no. A gene without classified based on the protein name to indicate the homology level such as same name, 'putative-' and '-like protein'. A gene without significant homology to any protein but with full-length cDNA or EST homology (covering almost the entire length of partial sequence) is classified as an 'unknown' protein. A gene predicted by two or more gene prediction programs is classified as a 'hypothetical' protein according to IRGSP standard. A gene predicted by a single gene prediction program is also classified as a probable 'hypothetical' protein and is included as a miscellaneous feature of the sequence. The orientation of the sequence is from SP6 to T7 of the PAC clone. This sequence of P0417G12 clone has an overlap with P0429G06 (DDBJ: AP003626) clone at 5' end and with P0691E09 (DDBJ: AP004781) at 3' end. Detailed information on overlap and assembly quality together with annotation of this entry is available at http://xgp.dna.affrc.go.jp/Genomeseq.html.
FEATURES	location/Qualifiers
source	1..163568
gene	/organism="Oryza sativa (japonica cultivar-group)"
gene	/mol_type="genomic DNA"
gene	/culivar="Nipponbare"
gene	/db_xref="caon:39947"
gene	/chromosome="6"
gene	/clone="P0417G12"
gene	2936..6301
gene	/gene="P0417G12.1"
gene	join(2936..3941,5348..6301)
gene	/gene="P0417G12.1"
gene	/note="supported by full-length cDNA(s) : AK071104"
gene	join(3060..3941,5348..5803)
gene	/gene="P0417G12.1"
gene	/note="contains full-length cDNA(s) : AK071104"
gene	plant zinc finger proteins"
gene	/codon_start=1
gene	/product="putative WIP1 protein"
gene	/protein_id="BAD35546.1"

repeat_unit	/db_xref="GI:51090943"
repeat_unit	/translation="MHTHTODTTLPLPLSLFIVVCLLHPI
repeat_unit	YVSSSMGDDVNTFLRGYHHHRHHHLHPPPP
repeat_unit	PSSPPIREALPLSLSTSTTHDDHHRRHQDHHHKKQCI
repeat_unit	DDKRESPSAAADDDQAEVVALHILPLSPSBDAAAKNK
repeat_unit	SSRMQVEEGEEDDEAAATAATPLPLGCAIGIGLTK
repeat_unit	QFSCPCVYKTFNRNNQMHHMGHSQYRKPSBELRTQPA
repeat_unit	IDHRARPPLDQFRLQTHYKKGIGKPFMCCKGKAFVRD
repeat_unit	GSDFKRRSLKDHIRAFGRGHAACIGISPDLDADDPSBDL
repeat_unit	10423..14032
repeat_unit	/gene="P0417G12.2"
repeat_unit	join(<10423..10459,13208..13706,13930..
repeat_unit	/gene="P0417G12.2"
repeat_unit	/note="start and end point are not identifi
repeat_unit	join(10423..10459,13208..13706,13930..1
repeat_unit	/gene="P0417G12.2"
repeat_unit	/note="predicted by GeneMark.hmm etc."
repeat_unit	/codon_start=1
repeat_unit	/product="hypothetical protein"
repeat_unit	/protein_id="BAD35547.1"
repeat_unit	/db_xref="GI:51090944"
repeat_unit	/translation="MAMHSGINKEKRGEDVCGTAGSGGHC
repeat_unit	GGRGGARRSGVAGETNARSGRSRSGSWPLAVELVETQ
repeat_unit	AAAELETLVAGVAAVYKRYNRKRRHSFGSSSTEDYD
repeat_unit	TTDADAASPPASPLSRCLRAPPPASAPRPPPAITDYLR
repeat_unit	complement(16882..23313)
repeat_unit	/gene="P0417G12.3"
repeat_unit	complement(16882..17848)
repeat_unit	/gene="P0417G12.3"
repeat_unit	/note="3' terminal repeat"
repeat_unit	/rpt_type="terminal
repeat_unit	complement(18655..22278)
repeat_unit	/gene="P0417G12.3"
repeat_unit	/note="probably inactive due to including :
repeat_unit	CDS
repeat_unit	pseudogene, Retrovirus-related Pol polyproi
repeat_unit	transposon TNT 1-94"
repeat_unit	/pseudo
repeat_unit	complement(22350..23313)
repeat_unit	/gene="P0417G12.3"
repeat_unit	/note="5' terminal repeat"
repeat_unit	/rpt_type="terminal
repeat_unit	23968..24174
repeat_unit	/gene="P0417G12.4"
repeat_unit	<23968..24174
repeat_unit	/gene="P0417G12.4"
repeat_unit	/note="start and end point are not identifi
repeat_unit	23968..24174
repeat_unit	/gene="P0417G12.4"
repeat_unit	/note="start and end point are not identifi
repeat_unit	23968..24174
repeat_unit	/gene="P0417G12.4"
repeat_unit	/note="predicted by GlimmerM etc."
repeat_unit	/codon_start=1
repeat_unit	/product="hypothetical protein"
repeat_unit	/protein_id="BAD35548.1"
repeat_unit	/db_xref="GI:51090945"
repeat_unit	/translation="MDASRAHSSAVGILNGERRAQQSRRG
repeat_unit	WKGDVGAGRWPAQPLFSTRGN"
repeat_unit	complement(32128..32409)
repeat_unit	/gene="P0417G12.5"
repeat_unit	complement(<32128..>32409)
repeat_unit	/gene="P0417G12.5"
repeat_unit	/note="start and end point are not identifi
repeat_unit	complement(32128..32409)
repeat_unit	/gene="P0417G12.5"
repeat_unit	/note="contains EST(s) : AU166246(E40406)"
repeat_unit	/codon_start=1
repeat_unit	/product="unknown protein"
repeat_unit	/protein_id="BAD35549.1"
repeat_unit	/db_xref="GI:51090946"
repeat_unit	/translation="MEMKKTAVAFVLAIAFAASATAVLASE
repeat_unit	APGAGGAAAGCAAGAAATSAAVVAAPPLAIVACS
repeat_unit	33317..33736
repeat_unit	/gene="P0417G12.6"
repeat_unit	<33317..>33736

/gene="P0417G12.6"
/note="start and end point are not identified"
3317..33736
/gene="P0417G12.6"
/note="predicted by Glimmer etc."
/codon_start=1
/product="hypothetical protein"
/protein_id="BADJ3550.1"
/translation="MKGGRHSPTRRHRTKATLCTRRTADRSVVCRLMRGCVGR
RRAAALHDARAGGAGGAGVODGDRPRVGAADPCRMAGVTCTVHTSSSSGPRV
GVAILNNLSASTCADRCITTCMCOVLESKTEAKLN"
complement(join(33966..34170,34708..34781))
/gene="P0417G12.7"
complement(join(33966..34170,34708..34781))
/gene="P0417G12.7"
/note="hypothetical ORF
predicted by GENSCAN
this category is not included in IRGSP standard"
complement(35929..36315)
/gene="P0417G12.8"
complement(35929..36315)
/gene="P0417G12.8"
/note="start and end point are not identified"
complement(35929..36315)
/gene="P0417G12.8"
/note="similar to Oryza sativa chromosome 3,
SUNBA0013D02.7"
/codon_start=1
/product="hypothetical protein"
/protein_id="BADJ3551.1"
/db_xref="GI:51090948"
/translation="MKRGQPRHADAFTPLWCHSLTTVALCFGRSRDWCSSATVP
NPRRHQASGGLPPSPVGNSSPDVEGVADFFGPTLDLMDLMDGDAQOK
GTLPASSSSSYAQRGAERAGRRP"
36448..36768
/gene="P0417G12.9"
36448..36768
/gene="P0417G12.9"
/note="hypothetical ORF
predicted by Glimmer
this category is not included in IRGSP standard"
37025..38599
/gene="P0417G12.10"
join(37025..37051,37279..37532,38222..38369,
38462..38599)
/gene="P0417G12.10"
/note="start and end point are not identified"

Alignment Scores:
Pred. No.: 1.71e-18 Length: 163568
Score: 419.50 Matches: 82
Percent Similarity: 60.00% Conservative: 26
Best Local Similarity: 45.56% Mismatches: 59
Query Match: 42.63% Indels: 13
DB: 15 Gaps: 2

US-10-716-089-20 (1-188) x AP003711 (1-163568)
QY 8 GlnGluThrThrSerGlnAlaProGluGlyProValMetCysValAsnLeuCysGlyPhe 27
DB 42873 AAGGAGACTGATGCCAGCAGCAGCAGCCGATCTTTCATCATTAATGCGGCTTC 42932
QY 28 PheGlySerGlnAlaThrMetGlyLeuCysSerIleCysValArgGluThrValMetGln 47
DB 42933 TTGGCAGTGGCGGATCATGATGCTCAAGTGCACACAGAGATGATTAATGAAG 42992
QY 48 AlaIysMetThrAlaLeuAlaGluGlnAlaThrGlnAlaAlaGlnAlaThrSerAlaThr 67
DB 42993 CAGGAGCAGCCAAAGCTGCA-----GCCTCCTCATTCGACAGCATGTCATGCTGT 43046
QY 68 AlaAlaAlaValGlnProProAlaProValIleGluThrIleLeuThrCysGluValGlu 87
DB 68 AlaAlaAlaValGlnProProAlaProValIleGluThrIleLeuThrCysGluValGlu 87

DB 43047 GATTCCGGAGAGAACCAATTATTGCTGTCACGTCGTAAGTAGCTGTTC 43106
QY 88 ArgThrMetIleValProHISGlnSerSerSerIleGlnAlaPheLeu 107
DB 43107 GTGAAGACGCTGT-----G 43133
QY 108 AlaAlaAlaProGlnAlaValIleSerSerIleAlaAlaProSerArgP 127
DB 43134 GAGATTGCTGCTCTTCAAGAGGAGGCTCAGCGTGAACCCCAAGGAGGAG 43195
QY 128 ArgCysGlySerCysArgIleArgValGlyLeuThrGlyPheIleCysVal 147
DB 43194 CGGTCTCCACTTGTCCGAAAGAGGTTGTCTTACCGGATTCACACTGC 43253
QY 148 LeuTYrCysAlaLeuHisArgTYrSerArgIleHisThrCysThrTYr 167
DB 43254 TTGTACTCGCAATGACACGCTAATCCGATAGCATGACTGCACTGTTC 43313
QY 168 AlaGluGlnGluAlaIleAlaIleValAlaPheProLeuValAlaGlu 187
DB 43314 GCTGCTAGGAGATGCTATTGCCAAGGCTAATCCGATGAGGCGGAG 43372

RESULT 14
AX654090 522 bp DNA linear -2003
DEFINITION Sequence 3960 from Patent WO03000898.
ACCESSION AX654090
VERSION AX654090.1 GI:29156904
KEYWORDS
SOURCE
ORGANISM
Oryza sativa
Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta
Spermatophyta; Magnoliophyta; Liliopsida; Poales; P
Euphorbiaceae; Oryzaceae; Oryza.
REFERENCE
1 Chang,H.S., Chen,W., Cooper,B., Glazebrook,J., Goff
Karagiri,F., Qian,S., Tao,Y., Whitman,S., Xie,Z., Z
Plant genes involved in defense against pathogens
Patent: WO 03000898-A 3960 03-JAN-2003;
SynGenea Participations AG (CH)
FEATURES
source location/Qualifiers
1..522
/organism="Oryza sativa"
/mol_type="unassigned DNA"
/db_xref="taxon:4530"

ORIGIN

Alignment Scores:
Pred. No.: 5.56e-21 Length: 522
Score: 416.50 Matches: 82
Percent Similarity: 57.29% Conservative: 28
Best Local Similarity: 42.71% Mismatches: 47
Query Match: 42.33% Indels: 35
DB: 6 Gaps: 3

US-10-716-089-20 (1-188) x AX654090 (1-522)
QY 8 GlnGluThrThrSerGlnAlaProGluGlyProValMetCysValAsnIle 27
DB 10 AAGAGGCTGCTGCTGCACACAGCCGAGGCGCCCATCTTCATCAATAT 69
QY 28 PheGlySerGlnAlaThrMetGlyLeuCysSerIleCysValArgGlu 47
DB 70 TTGGCAGTGGCGGACCATTAACATGCTCCCAAGTGCACAGAGAG 129
QY 48 AlaIysMetThrAlaLeuAlaGluGlnAlaThrGlnAla----- 60
DB 130 CAGGAGCAGCCAAAGCTTGTCTCTCTCCATTCGATGACATTCGATTC 189
QY 61 -----AlaGlnAlaThrSerAlaThrAlaAlaValG 75
DB 190 GGAAGAGCAGCATATTGTTGTCGACAGTGTAGCACGCGCGTGC 237

Oy	76	ProValHtAGluThrThylLeuThrCysGluValGluArgThrMetIleValProHtAGln	95
Db	238	-----GCTCAGGTCGAGCGGAGACCGCTGTT-----	264
Oy	96	SerSerSerTyrgInGlnAspLeuValThrProAlaAlaAlaProGlnAlaVallys	115
Db	265	-----GTGCAGCGCTTACCAGTGTCCGGGACCGGACGACGAGAG	300
Oy	116	SerSerIleAlaAlaProSerArgProGluProAlaAspGlySerCysArgArgysArg	135
Db	301	GTTCGTGTATGTCCTCCAGAGTCAAGAGCCGGAACCGGTGGCTTACCTGTATGGAAGG	360
Oy	136	ValGlyLeuThrGlyPheLeuCysArgCysGlyAsnLeuTyrcysAlaLeuHtIstArgTy	155
Db	361	GTTCGGCTGACCGGATTCACATCGCGGTGCGGTAAACATGTACTGTGCGTTGACCGCTAC	420
Oy	156	SerAspIlyHtIstHrThyrThyrAspTyrylyAspAlaAlaGlyGlnGlnAlaIleAllys	175
Db	421	TCCGCAAGCATGATGATGCGACGATTCACCGACTCCGCGCTAGGAGATGCCATGCCCAAG	480
Oy	176	AlaAsnProLeuValAlaAlaGluIysValIlys	187
Db	481	GCCAAACCAAGTGTTGAAGCTGAAGCTCCACAAAG	516
RESULT 15	AK061911	991 bp mRNA linear	PLN 24-JUL-2003
LOCUS	Oryza sativa (japonica cultivar-group)	cdna clone:001-042-A06, full	
DEFINITION	Insert sequence.		
ACCESSION	AK061911		
VERSION	AK061911.1	GI:32971929	
KEYWORDS	FLI CDNA; oligo-capping.		
SOURCE	Oryza sativa (japonica cultivar-group)		
ORGANISM	Oryza sativa (japonica cultivar-group)		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.		
REFERENCE	1 The Rice Full-Length cDNA Consortium, National Institute of Agricultural Sciences Rice Full-Length cDNA Project Team; Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yasaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, M., Suzuki, K., Li, C., Ohtsuka, K., Shishiki, T., Ohneda, E., Yahagi, M., Suzuki, K., Li, C., Ohtsuka, K., Shishiki, T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group; Otsu, Y., Murakami, K., Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y., Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Naitakawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niihara, J., Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Mura, J., Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN, Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Otsu, N., Ota, Y., Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Yoshino, M., and Hayashizaki, Y.		
TITLE	Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice		
JOURNAL	Science 301 (5631), 376-379 (2003)		
PUBMED	12869764		
REFERENCE	2 (bases 1 to 991)		
AUTHORS	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoaka, T., Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K., Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M., Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Mura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, K., Namiki, T., Naitakawa, R., Niihara, J., Nishikawa, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuka, K., Oka, M., Ooka, H.,		

Otsu, N., Ota, Y., Otsu, Y., Ryu, R., Satoh, H., Saka
Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K.,
Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., S
Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagam
Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka,
Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yaka
Yamada, H., Yamamoto, M., Yasunishi, A., Yasaki, J., Yc
Yoshimura, A.

TITLE
Direct Submission

JOURNAL
Submitted (05-DEC-2001) Shoshi Kikuchi, National In
Agrobiological Sciences, Department of Molecular Ge
Laboratory of Gene Expression; 2-1-2 Kannondai, Tsu
305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp,
Tel:81-29-838-7007, Fax:81-29-838-7007)

COMMENT
This clone is one of the 28K full-length cDNA clone:
rice.

URL: <http://cdna01.dna.affrc.go.jp/cDNA/>

NIAS Rice Full-length cDNA Project Team: Kikuchi, S.,
Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Ya
Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K.
Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuka, K.,
Yamamoto, M.

FAIS Genome Sequencing & Analysis Group: Otsu, Y.,
Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Ko
Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, I
Mizuno, K., Naitakawa, R., Niihara, J., Oka, M., Ryu, R.,
Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q.,
Yoshimura, A., Matsubara, K., and Murakami, K.

**Genome Exploration Research Group in Riken Genomic
and Genome Science Laboratory in Riken**: Adachi, J.,
Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Han
Hara, A., Hashizume, W., Hayatsu, N., Hotta, I.,
Hiroaka, T., Hori, F., Iida, J., Imamura, K., Imotani, K.
Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J.,
Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H.
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Mu
Nakamura, M., Nishii, K., Nomura, K., Numasaki, R., Ohn
Ota, Y., Satoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shira
Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A.,
Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., W
Yasunishi, A., and Hayashizaki, Y.

FEATURES	source		
ORIGIN	1..991		
Alignment Scores:	Location/Qualifiers		
Pred. No.:	1..116-20		
Score:	416.50		
Percent Similarity:	57.29%		
Best Local Similarity:	42.71%		
Query Match:	42.33%		
DB:	15	Gaps:	3
US-10-716-089-20 (1-188) x AK061911 (1-991)			
Oy	8	GlnGluThrThrSerGlnAlaProGluGlyProValMetCysIysAsnL	27
Db	153	AAGAGGCTGCTGCTCCAGCAGCCGAGGCGCCATCTTCATCAATAA	212
Oy	28	PheGlySerGlnAlaThrMetCylLeuCysSerIysCysTyraGluT	47
Db	213	TTTGGCAGCGCGGCGACCATGTAACATGTCTCCAAAGTCCCAAGAGA	272
Oy	48	AlaIysMetThrAlaLeuAlaGluGlnAlaThrGlnAla-----	60
Db	273	GAGGAGCAGGCCAAGCTTGTGCTCTCCATCATGATGATTCATCAATG	332


```
QY 61 -----AlaGlnAlaThrSerAlaThrAlaAlaValGlnProProAla 75
Db 333 GGAAGAGCATATTGTTGCTGCTGAGCAGCGCGGTGGCGTG----- 380
QY 76 ProValHLeGluThrLeuThrCysGluValGluArgThrMetIleValProHLeGln 95
Db 381 -----GCTCAGGTGAGGCGAAGACGCTCGTT----- 407
QY 96 SerSerSerTyrgInGlnApleuValThrProAlaAlaAlaProGlnAlaValIys 115
Db 408 -----GTGCAGCCTTACCGATGTCGCGGCGACAGCGAGAG 443
QY 116 SerSerIleAlaAlaProSerArgProGluProAlaArgCysGlySerCysArgIysArg 135
Db 444 GTTGGCTGATGCTCCCAAGGTCAAGAGGCGCAACCGGTGCGCTACCTGTAGGAAGAG 503
QY 136 ValGlyLeuThrGlyPheIysCysArgCysGlyAsnLeuTyrcysAlaLeuHIsArgTyr 155
Db 504 GTTGGGCTGACCGGATTCACTGCCGCTCGGTAACTACTGTGCGTTGCACCGCTAC 563
QY 156 SerAspLysHIsThrCysThrTyraerTyrLysAlaAlaGlyGlnGluAlaIleAlaIys 175
Db 564 TCCGACAAGCATGAATGCGACGTACCGGACTGCGGCTAGGGATGCCATCGCCNAG 623
QY 176 AlaAsnProLeuValValAlaGluLysValValIys 187
Db 624 GCCAACCCAGTGTGAGGCTGAGAACTCGACNAG 659
```

Search completed: December 8, 2005, 17:40:21
Job time : 3337 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: December 8, 2005, 16:32:11 / Search time 472 Seconds
(without alignments)
2654.561 Million cell updates/sec

Title: US-10-716-089-20

Sequence: 1 MATERVSETTSQAEPGPM.....GQEAIAKANPLVAKVKYKF 188

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 4996997 seqs, 332346308 residues
Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame_plus.p2n.model -DEV=xlp
-O=/cgn2.1/USPTO_spool_p/US10716089/runac_05122005_094343_13131/app_query.fasta_1.327
-DB=N Geneseq -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsu62 -TRANS=human40.cdi -LIST=45
-OUTALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10716089@cgn2.1 1096@runac_05122005_094343_13131 -NCPU=6 -ICPU=3
-NO_WMAP -LARGESQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOC
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_21.*
1: geneseqn1980s:*
2: geneseqn1980s:*
3: geneseqn2000s:*
4: geneseqn2001s:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*
14: geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	984	100.0	804	6	ABST0622 Transcript
2	984	100.0	804	6	AAS95832 CDNA enco
3	957	97.3	1531	6	ABST0614 Transcript
4	957	97.3	1531	6	AAS95824 Partial c

5	425.5	43.2	945	13	ADX51406	Adx
6	419.5	42.6	1000	14	ABE65909	Adx
7	419.5	42.6	1021	14	ABE65932	Adx
8	416.5	42.3	519	11	ACL30227	Adx
9	416.5	42.3	522	8	ADA70637	Adx
10	416.5	42.3	522	11	ACL27951	Adx
11	416.5	42.3	1032	13	ADX52834	Adx
12	415	42.2	522	11	ACL34078	Adx
13	415	42.2	736	13	ADX47307	Adx
14	415	42.2	776	13	ADX53784	Adx
15	415	42.2	821	13	ADX47342	Adx
16	415	42.2	826	13	ADX46752	Adx
17	415	42.2	841	13	ADX50951	Adx
18	415	42.2	846	13	ADX50987	Adx
19	415	42.2	853	13	ADX54045	Adx
20	415	42.2	863	13	ADX52432	Adx
21	415	42.2	863	13	ADX53800	Adx
22	415	42.2	868	13	ADX33018	Adx
23	415	42.2	868	13	ADX54028	Adx
24	415	42.2	872	13	ADX51702	Adx
25	415	42.2	872	13	ADX53116	Adx
26	415	42.2	873	13	ADX53121	Adx
27	415	42.2	876	13	ADX52889	Adx
28	415	42.2	886	13	ADX12369	Adx
29	415	42.2	894	13	ADX48800	Adx
30	415	42.2	898	13	ADX6184	Adx
31	415	42.2	899	13	ADX50721	Adx
32	415	42.2	902	13	ADX52776	Adx
33	415	42.2	904	13	ADX47138	Adx
34	415	42.2	918	13	ADX61202	Adx
35	415	42.2	920	13	ADX47807	Adx
36	415	42.2	924	13	ADX60812	Adx
37	415	42.2	926	13	ADX52877	Adx
38	415	42.2	931	13	ADX52869	Adx
39	415	42.2	938	13	ADX53497	Adx
40	415	42.2	932	13	ADX51034	Adx
41	415	42.2	942	13	ADX53778	Adx
42	415	42.2	943	13	ADX5613	Adx
43	415	42.2	948	13	ADX10810	Adx
44	415	42.2	953	13	ADX51073	Adx
45	415	42.2	956	13	ADX53049	Adx

ALIGNMENTS

RESULT 1	
ABST0622	
ID	ABST0622 standard; cDNA; 804 BP.
XX	
AC	ABST0622;
DT	28-NOV-2002 (first entry)
XX	
DE	Transcription factor stress-related protein (TFSRP), cDNA
KW	Transcription factor stress-related protein; TFSP; MYB-1
KW	AP2 similar-2 protein; APS-2; zinc-finger factor-2; ZF-2;
KW	ZF-4 protein; ZF-5 protein; CAAT-Box binding factor; proxi
KW	Sigma factor like protein; SFL-1; transgenic; plant; dros
KW	environmental stress; Physcomitrella patens; high salinity
KW	climate; fungus; gene; ss.
OS	Physcomitrella patens.
XX	
PN	US2002102695-A1.
XX	
PD	01-AUG-2002.
XX	
PF	06-APR-2001; 2001US-00828303.
XX	
PR	07-APR-2000; 2000US-0196001P.
XX	
PA	(SILV/) SILVA O D C E.

PA (BOHNERT H J.
PA (THIELEN N V.
XX (CHEN)/ CHEN R.
XX
XX Silva ODCR, Bohnert HJ, Thiele N, Chen R,
XX
XX WPI, 2002-690614/74.
XX
XX P-PsDB; ABG93809.

PT Novel transcription factor stress-related protein such as AP2 similar-2
PT protein, zinc-finger protein, or CAAT-Box binding factor protein-3,
PT useful for modifying stress tolerance of plant.

PS Claim 3; Fig 2D; 104pp; English.

XX The invention relates to an isolated transcription factor stress-related
XX protein (TFSRP) (I) selected from AP2 similar-2 protein (AP2-2), zinc-
XX finger factor-2 (ZF-2) protein, ZF-3 protein, ZF-4 protein, ZF-5 protein,
XX MYB-1 protein, CAAT-Box binding factor protein-3 (CAAT-3), Sigma factor
XX like protein (SFL-1) and their orthologues. Also described is: (1) an
XX isolated TFSRP coding nucleic acid (II) coding for (I); (2) a transgenic
XX plant cell (III) transformed by (II), where the expression of (II) in the
XX plant cell results in increased tolerance to an environmental stress as
XX compared to a wild-type variety of the plant cell; (3) a transgenic plant
XX (IV) comprising (III); (4) a seed (V) produced by (IV), where the seed is
XX true breeding for an increased tolerance to environmental stress as
XX compared to a wild-type variety of the plant cell; (5) an agricultural
XX product (VI) produced by (IV) or (V); and (6) an isolated recombinant
XX expression vector (VII) comprising (II), where expression of (VII) in a
XX host cell results in increased tolerance to environmental stress as
XX compared to a wild-type variety of the host cell. (II) is useful for
XX modifying (increasing or decreasing) stress tolerance of a plant
XX (transgenic or not transgenic), by modifying the expression of TFSRP in
XX the plant. The plant is transformed with a promoter that directs
XX expression of the TFSRP. The promoter is tissue specific and is
XX developmentally regulated. TFSRP expression is modified by administration
XX of an antisense molecule that inhibits expression of TFSRP. (VII) is
XX useful for producing a transgenic plant containing (II), where expression
XX of the nucleic acid in the plant results in increased tolerance to
XX environmental stress as compared to a wild-type variety of the plant, by
XX transforming a plant cell with (VII) comprising the nucleic acid,
XX generating from the plant cell a transgenic plant with an increased
XX tolerance to environmental stress as compared to a wild-type variety of
XX the plant. (I) is useful for conferring stress tolerance such as drought,
XX cold and/or salt tolerance to plants. (I), (II), (III) or (VII) is useful
XX for identifying Physcomitrella patens and related organisms, for mapping
XX of genomes of organisms related to P.patens, for identifying and
XX localizing P.patens sequences of interest, for evolutionary studies, for
XX determining TFSRP regions required for function, for modulating TFSRP
XX activity, for modulating metabolism of one or more cell functions, for
XX modulating transmembrane transport of one or more compounds, and for
XX modulating stress resistance. (I) is useful for reducing stress tolerance
XX response plants or more particularly, in the transcription of a protein
XX involved in a stress tolerance response in a P.patens plant. (II) is
XX useful for transforming plants and thus inducing tolerance to stresses
XX such as drought, high salinity and cold, for identifying the presence of
XX P.patens or a related organism in a mixed population of microorganisms,
XX serve as marker for specific regions of the genome for mapping the genome
XX and for the functional studies of P.patens proteins. (II) is useful as
XX reference points for mapping the moss genomes, or of genomes of related
XX organisms, for evolutionary and protein structural studies, for
XX generating knockout mutation in the genomes of various organisms such as
XX bacteria, mammalian cells, yeast cells and plant cells, useful for
XX evaluating their ability or capacity to tolerate various stress
XX conditions and the effect on the phenotype and/or genotype of the
XX mutation. (II) is useful as marker for the construction of a genomic map
XX in related mosses. (I) or (II) is useful for generating mutated TFSRP
XX plants, fungi or other microorganisms expressing mutated TFSRP nucleic
XX acid and protein molecules for improving stress tolerance. AB570611-
XX AB570681 represent P. patens TFSRP coding sequences and PCR primers of
XX the invention

Sequence 804 BP; 186 A; 214 C; 205 G; 199 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.98e-77 Length: 804
Score: 984.00 Matches: 188
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-716-089-20 (1-188) x AB570622 (1-804)

QY 1 MetAlaThrgluArgValSerGlnGluThrThrsSerGlnAlaProGluG.
DB 134 ATGGCCACCGAGCGTGTGTCTCAGGAGACGACCTCGCAGGCCCTGAGG
QY 21 CysAlaSerLeuCyGlyGlyPhePheGlySerGlnAlaThrMetGlyLeuCy
DB 194 TGCAGGAACCTTTCGGCTTCTTGGCAGCCAGCTACCATGGGGTGTCTC
QY 41 TyrArgGluThrValMetGlnAlaValMetThrAlaLeuAlaGluGlnA
DB 254 TACCGAGACAGTCATGCAAGCAGACGATGACGCTTTAGCTGAGCAG
QY 61 AlaGlnAlaThrSerAlaThrAlaAlaAlaValGlnProAlaProV
DB 314 GCTCAGGCGACATCTGCACAGCTGCTGTTTACGCCCTCCTCTG
QY 81 LysLeuThrCyGlyValGluValGluThrMetLeuValProHisGlnSerS
DB 374 AAGCTCACATCGCAGGTTGAGAGAACATGATGTTGCGCATCAATCT
QY 101 GlnPheLeuValThrProAlaAlaAlaAlaProGlnAlaValSerS
DB 434 CAAAGCTGGTTACCCCGCTGCAGCTGCCCTTACGAGCAGGAACTCT
QY 121 ProSerArgProGluProAsnArgCyGlySerCyArglyValArgValG
DB 494 CCGCTTGACCCGAGCCCAATGATGCGGATCTTGACGAGGACCGTGTG
QY 141 PheLysCyArgCyGlyValSerLeuThrCyValAlaLeuHisAlaGlyT
DB 554 TTTAAGTCTCCCTGTGGCAACCTTACTGCGCTTACTCGTACTCGG
QY 161 CysThrTyrAspTyrLysAlaAlaGlyGlnGlnAlaAlaAlaValAlaA
DB 614 TGCATATGACTCAAGCCGCGAGGCGAGAGCGATTCGGAAGCTA
QY 181 ValAlaGluValValLysPhe 188
DB 674 GTGGCCGAGAAAGTTGTCAAAGTT 697

RESULT 2
AAS95832 standard; cDNA; 804 BP.
AAS95832

AAS95832;

26-FEB-2002 (first entry)

cDNA encoding Zinc finger-4 (ZF-4).

Transcription factor stress-related protein; AP2 similar-2

AP2-2; zinc-finger factor protein; ZF-2; ZF-3; ZF-4; ZF-5;

MYB-1 MYB-1 protein; CAAT-box binding factor protein-3; CA

Sigma factor like protein; SFL-1; environmental stress; mo

Physcomitrella patens.

WO200177311-A2.

18-OCT-2001.

PF 06-APR-2001; 2001WO-US011393.
 XX 07-APR-2000; 2000US-0196001P.
 XX (BADI) BASF PLANT SCI GMBH.
 XX Da Costa Silva O, Bohnert HJ, Van Thielén N, Chen R;
 XX MPI 2002-049146/06.
 DR P-PSDB; AAU71800.
 XX
 PT New polypeptide, useful for identification of Physcomitrella patens, and
 PT for modulating stress resistance of a plant, comprises an isolated
 PT transcription factor stress-related protein derived from Physcomitrella
 PT patens.
 XX
 XX Claim 14; Fig 2D; 129p; English.
 PS
 XX
 CC The invention relates to an isolated transcription factor stress-related
 CC protein (TFSRP) (I) selected from an AP2 similar-2 protein (APS-2), a
 CC zinc-finger factor-2 protein (ZF-2), ZF-3 protein, ZF-4 protein, ZF-5
 CC protein, a MYB-1 protein (MYB-1), a CAAT-box binding factor protein-3
 CC (CAF-3), a Sigma factor like protein (SFL-1) and its orthologues. Also
 CC described is an isolated TFSRP coding nucleic acid (II) which codes for
 CC (I); and an isolated recombinant expression vector (III) comprising (II),
 CC where expression of (III) in a host cell results in increased tolerance
 CC to environmental stress as compared to a wild type variety of the host
 CC cell. (I) or (II) is useful as markers for specific regions of the TFSRP
 CC genome. (I) or (II) is useful for identification of Physcomitrella patens
 CC and related organisms, mapping of genomes of organisms related to P.
 CC patens, identification studies, determination of P. patens sequence of
 CC interest, evolutionary studies, determination of TFSRP regions required
 CC for function, modulation of a TFSRP activity, modulation of the
 CC metabolism of one or more cell functions, modulation of the transmembrane
 CC transport of one or more compounds and modulation of stress resistance.
 CC (II) is useful for generating probes and primers for identifying and/or
 CC cloning TFSRP homologues in other cell types and other organisms, as well
 CC as TFSRP homologues from mosses and related species, and for evolutionary
 CC and protein structural studies. AAS95821-AAS95891 represent P. patens
 CC TFSRP coding sequences and PCR primers of the invention
 CC
 SQ Sequence 804 BP; 186 A; 214 C; 205 G; 199 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1,986-77 Length: 804
 Score: 984.00 Matches: 188
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0
 US-10-716-089-20 (1-188) x AAS95832 (1-804)
 QY 1 MetAlaThrGluArgValSerGlnGluThrThrSerGlnAlaProGluGlyProValMet 20
 DB 134 ATGGCCACCGAGCGTGTGCTTCAGAGAGACGACCTCGCAGGCCCTGAGGGTCAAGTATAG 193
 QY 21 CysLysAsnLeuCysGlyPhePheGlySerGlnAlaThrMetGlyLeuCysSerLysCys 40
 DB 194 TGCAGAACCTTGGCGGCTTCTTCGCGCAGCCAAAGTACATGAGGGTGGTGGCTCGAAGTGC 253
 QY 41 TyrArgGluThrValMetGlnAlaLysMetThrAlaLeuAlaGluGlnAlaThrGlnAla 60
 DB 254 TACCAGAGAGACGTATGCAACGAGAGAGAGCGCTTAAAGTACGACCAAGCCACTCAGGCT 313
 QY 61 AlaGlnAlaThrSerSerAlaThrAlaAlaValAlaGlnProProAlaProValHisGlnThr 80
 DB 314 GCTCAGGGGCACTTGTCCACACCTGCTGCTTTCAGCCCCCGCTCCTGTACATAGAC 373
 QY 81 LysLeuThrCysGluValAlaGluArgThrMetCileValProHisGlnSerSerSerTyrgln 100
 DB 374 AAGTCACATCGGAGGTTCAGAGAACCAATGATGTGGCCGATCAATCTTCCAGCTATCA 433

QY 101 GlnApeLewValThrProAlaAlaAlaProGlnAlaValLysSer 120
 DB 434 CAAGACCTGTTACCCCGCTGCAGCTGCCCTCGAGGAGTGAATGCT 493
 QY 121 ProSerArgProGluProAsnArgCysGlySerCysArgLysArgVal 140
 DB 494 CCTTACAGCCGAGACCCCAATGATGATGCGGATCTTCAGAGACGTTTC 553
 QY 141 PheLysCysArgCysGlyAsnLeuTyrCysAlaLeuHisArgTyrSer 160
 DB 554 TTTAAGTGTGCGTGTGGCAACCTCTACGCTTACATCGGTACTCGG 613
 QY 161 CysThrTyrApeTyrLysAlaAlaGluGlnGluValAlaAlaVal 180
 DB 614 TGCACATATGACTACAAAGCCGACGGGACGAGAAAGCAATTCGAAAGCTA 673
 QY 181 ValAlaGluLysValValLysPhe 188
 DB 674 GTGGCCGAGAGGTTGTCAAGTTT 697
 RESULT 3
 ID ABS70614 standard; cDNA; 1531 BP.
 XX ABS70614;
 AC 28-NOV-2002 (first entry)
 XX
 DT Transcription factor stress-related protein (TFSRP), parti
 XX
 DE Transcription factor stress-related protein (TFSRP), parti
 XX
 KW Transcription factor stress-related protein; TFSP; MYB-1
 KW AP2 similar-2 protein; APS-2; zinc-finger factor-2; ZF-2;
 KW ZF-4 protein; ZF-5 protein; CAAT-box binding factor; protei
 KW Sigma factor like protein; SFL-1; transgenic; plant; droug
 KW environmental stress; Physcomitrella patens; high salinity
 KW ciliate; fungus; gene; ss.
 XX
 OS Physcomitrella patens.
 XX
 PN US2002102695-A1.
 XX
 PD 01-AUG-2002.
 XX
 PF 06-APR-2001; 2001US-00828303.
 XX
 PR 07-APR-2000; 2000US-0196001P.
 XX
 PA (SILV/) SILVA O D C E.
 PA (BOHN/) BOHNERT H J.
 PA (THIE/) THIELEN N V.
 PA (CHEN/) CHEN R.
 XX
 PI Silva ODE, Bohnert HJ, Thielén NV, Chen R;
 XX
 DR MPI 2002-690614/74.
 XX
 PT Novel transcription factor stress-related protein such as
 PT protein, zinc-finger protein, or CAAT-Box binding factor p
 PT useful for modifying stress tolerance of plant.
 XX
 PS Example 5; Fig 1D; 104p; English.
 XX
 CC The invention relates to an isolated transcription factor
 CC protein (TFSRP) (I) selected from AP2 similar-2 protein (A
 CC finger factor-2 (ZF-2) protein, ZF-3 protein, ZF-4 protein
 CC MYB-1 protein, CAAT-box binding factor protein-3 (CAF-3),
 CC like protein (SFL-1) and their orthologues. Also described
 CC isolated TFSRP coding nucleic acid (II) coding for (I); (I)
 CC plant cell (III) transformed by (II), where the expression
 CC plant cell results in increased tolerance to an environmen
 CC compared to a wild-type variety of the plant cell; (3) a t
 CC (IV) comprising (III); (4) a seed (V) produced by (IV), wh
 CC true breeding for an increased tolerance to environmental

CC compared to a wild-type variety of the plant cell; (5) an agricultural
CC product (VI) produced by (IV) or (V); and (6) an isolated recombinant
CC expression vector (VII) comprising (II), where expression of (VII) in a
CC host cell results in increased tolerance to environmental stress as
CC compared to a wild-type variety of the host cell. (II) is useful for
CC modifying (increasing or decreasing) stress tolerance of a plant
CC (transgenic or not transgenic), by modifying the expression of TFSRP in
CC the plant. The plant is transformed with a promoter that directs
CC expression of the TFSRP. The promoter is tissue specific and is
CC developmentally regulated. TFSRP expression is modified by administration
CC of an antisense molecule that inhibits expression of TFSRP. (VII) is
CC useful for producing a transgenic plant containing (II), where expression
CC of the nucleic acid in the plant results in increased tolerance to
CC environmental stress as compared to a wild-type variety of the plant, by
CC transforming a plant cell with (VII) comprising the nucleic acid,
CC generating from the plant cell a transgenic plant with an increased
CC tolerance to environmental stress as compared to a wild-type variety of
CC the plant. (I) is useful for conferring stress tolerance such as drought,
CC cold and/or salt tolerance to plants. (I), (II), (III) or (VII) is useful
CC for identifying Physcomitrella patens and related organisms, for mapping
CC of genomes of organisms related to P.patens, for identifying and
CC localizing P.patens sequences of interest, for evolutionary studies, for
CC determining TFSRP regions required for function, for modulating TFSRP
CC activity, for modulating metabolism of one or more cell functions, for
CC modulating transmembrane transport of one or more compounds, and for
CC modulating stress resistance. (I) is useful for reducing stress tolerance
CC response plants or more particularly, in the transcription of a protein
CC involved in a stress tolerance response in a P.patens plant. (II) is
CC useful for transforming plants and thus inducing tolerance to stresses
CC such as drought, high salinity and cold, for identifying the presence of
CC P.patens or a related organism in a mixed population of microorganism,
CC serve as marker for specific regions of the genome for mapping the genome
CC and for the functional studies of P.patens proteins. (II) is useful as
CC reference points for mapping the moss genomes, or of genomes of related
CC organisms, for evolutionary and protein structural studies, for
CC generating knockout mutation in the genomes of various organisms such as
CC bacteria, mammalian cells, yeast cells and plant cells, useful for
CC evaluating their ability or capacity to tolerate various stress
CC conditions and the effect on the phenotype and/or genotype of the
CC mutation. (II) is useful as marker for the construction of a genomic map
CC in related mosses. (I) or (II) is useful for generating algae, ciliates,
CC plants, fungi or other microorganisms expressing mutated TFSRP nucleic
CC acid and protein molecules for improving stress tolerance. ABS70611-
CC ABS70681 represent P. patens TFSRP coding sequences and PCR primers of
CC the invention
XX
XX
SQ Sequence 1531 BP, 313 A; 399 C; 393 G; 426 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1,046-74 Length: 1531
Score: 957.00 Matches: 187
Percent Similarity: 99.47% Conservative: 0
Best Local Similarity: 99.47% Mismatches: 1
Query Match: 97.26% Indels: 1
DB: 6 Gaps: 0

US-10-716-089-20 (1-188) x ABS70614 (1-1531)

QY 1 MetAlaThrGluAlaValSerGlnGluThrSerGlnAlaProGlnGlyProValMet 20
DB 655 ATGGCCACCGAGCGGTGTCTCAGAGACGACCTCGCAGGCCCCGAGGGTCCAGATTATG 714
QY 21 CybelysaantleuCybGlyPhePheGlySerGlnAlaThrMetGlyLeuCybSerlyGys 40
DB 715 TGCAGAACCTTGGCTTCTTCGCGACGCCAAGCTACCATGGGATTGGCTCGAAGTGC 774
QY 41 TyrArgGluThrValMetGlnAlaIleMetThrAlaLeuAlaGlnGlnAlaThrGlnAla 60
DB 775 TACCGAGACAGCTATGTCAGCGC-AAAGTACAGCGCTTTAGCTGACACAGCCACTAGGCT 833
QY 61 AlaGlnAlaThrSerAlaThrAlaAlaValAlaGlnProProAlaProValHisGlnThr 80
DB * 834 GCTCAGCGGACATCTGCACACAGCTGCTGCTTTCAGCCGCCCGCTCCTGATCATGAGACC 893

QY 81 LysLeuThrCybGlyValGluAlaThrMetIleValProHisGlnSerS 100
DB 894 AAGTCCATCTCGAGGTGTGAGAAACATGATGTGGCCGATCAATCTT 953
QY 101 GlnAspLeuValThrProAlaAlaAlaAlaProGlnAlaValSerS 120
DB 954 CAAGACCTGGTATCCCGCGCTGCGAGCTGCCCTCAGGCGAGTGAAGTCTT 1013
QY 121 ProSerArgProGluProAlaMetCybGlySerCybArglyValG 140
DB 1014 CCCCTGACCCCGAGCCCAATCGATGCGGATCTTGACGAGAACCGTGTG 1073
QY 141 PheLysCybArgCybGlyAsnLeuThrCybAlaLeuHisArgTyrSer 160
DB 1074 TTTAAGTGTCCCTGTGGCAACCTTACTGCGCTTATCTGCTACTCGG 1133
QY 161 CybThrTyrAspTyrLysAlaAlaGlnGlnAlaIleAlaValS 180
DB 1134 TGCAATATGACTCAAAAGCCGAGGCGAGGACGATTCGAAAGCTA 1193
QY 181 ValAlaGluLysValValLysPhe 188
DB 1194 GTGGCCGAGAAAGTGTTCAACTTT 1217

RESULT 4

IDS AAS95824 standard; cDNA; 1531 BP.
XX AAS95824;
AC AAS95824;
XX 26-FEB-2002 (first entry)
DT Partial cDNA encoding zinc finger-4 (ZF-4).
XX
DE Transcription factor stress-related protein; AP2 similar-2
XX AP2-2; zinc-finger factor protein; ZF-2; ZF-3; ZF-4; ZF-5;
XX MYB-1 MYB-1 protein; CAAT-box binding factor protein-3; CAI
XX Sigma factor like protein; SFL-1; environmental stress; mo
XX 85.
OS Physcomitrella patens.
XX
XX MO200177311-A2.
XX 18-OCT-2001.
PD 06-APR-2001; 2001WO-US011393.
XX 07-APR-2000; 2000US-0196001P.
PR (BADI) BASF PLANT SCI GMBH.
XX Da Costa Silva O, Bohnert HJ, Van Thieleen N, Chen R;
XX WPI, 2002-049146/06.
DR
XX
XX New polypeptide, useful for identification of Physcomitrel.
PT For modulating stress resistance of a plant, comprises an
PT transcription factor stress-related protein derived from Pl
PT patens.
XX
XX Example 5; Fig 1D; 129pp; English.
PS
XX The invention relates to an isolated transcription factor
XX protein (TFSRP) (I) selected from an AP2 similar-2 protein
XX zinc-finger factor-2 protein (ZF-2), ZF-3 protein, ZF-4 pr
XX protein, a MYB-1 protein (MYB-1), a CAAT-box binding factor
XX (CABF-3), a Sigma factor like protein (SFL-1) and its orth
XX described is an isolated TFSRP coding nucleic acid (II) wh
XX (I); and an isolated recombinant expression vector (III) c
XX where expression of (III) in a host cell results in increa
XX to environmental stress as compared to a wild type variety

CC cell. (I) or (II) is useful as markers for specific regions of the TFSRP
CC genome. (I) or (II) is useful for identification of phycometrella patens
CC and related organisms, mapping of genomes of organisms related to P.
CC patens, identification and localization of P. patens sequence of
CC interest, evolutionary studies, determination of TFSRP regions required
CC for function, modulation of a TFSRP activity, modulation of the
CC metabolism of one or more cell functions, modulation of the transmembrane
CC transport of one or more compounds and modulation of stress resistance.
CC (II) is useful for generating probes and primers for identifying and/or
CC cloning TFSRP homologues in other cell types and other organisms, as well
CC as TFSRP homologues from mosses and related species, and for evolutionary
CC and protein structural studies. AAS95821-AAS95891 represent P. patens
CC TFSRP coding sequences and PCR primers of the invention
XX
SQ Sequence 1531 BP; 313 A; 399 C; 393 G; 426 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 1,04e-74 Length: 1531
Score: 957.00 Matches: 187
Percent Similarity: 99.47% Conservative: 0
Best Local Similarity: 99.47% Mismatches: 1
Query Match: 97.26% Indels: 1
DB: Gaps: 0
US-10-716-089-20 (1-188) x AAS95824 (1-1531)
QY 1 MetAlaThrGluArgValSerGlnGluThrSerGlnAlaProGlnGlyProValMet 20
DB 655 ATGGCCACCGAGCGGTGTCTCAGGAGAGACCTCGCAGGCCCTGAGGGTCCAGTTATG 714
QY 21 CysIlyAsnIleuCyGlyPhePheGlySerGlnAlaThrMetGlyIleuCySerIlyCys 40
DB 715 TCCAGAACCTTGGCGCTTCTTCCGACCCAGCTACCATGGGTGTGCTCGAAGTGC 774
QY 41 TyrArgGluThrValMetGlnAlaIlyMetThrAlaIleuAlaGluGlnAlaThrGlnAla 60
DB 775 TACCGAGAGACAGTATGACAGCG -AAGATGACGGGCTTACGAGAGACCACTCAGGCT 833
QY 61 AlaGlnAlaThrSerAlaThrAlaAlaAlaValGlnProAlaProValHisGluThr 80
DB 834 GCTCAGCGCAGCTCTGCCACAGCTCTGCTTACAGCCCCCGCTCTGTACATGAGACC 893
QY 81 LysIleuThrCysGluValaGluArgThrMetIleValProHisGlnSerSerSerTyrGln 100
DB 894 AAGCTCACATGGAGAGTTGAGAAACAATGATTTGCCGCAATCAATCTTCCAGCTATGCA 953
QY 101 GlnAsePleuValThrProAlaAlaAlaAlaProGlnAlaValIlySerSerIleAlaAla 120
DB 954 CAAAGACTGGTTACCCCGCTGACGTGCCCTCAGGCGATGAATCTCTTATCCGAGCT 1013
QY 121 ProSerArgProGlnProAlaAlaIlyGlySerCyArgIlyArgValGlyIleuThrGly 140
DB 1014 CCCTCTAAGCCCGAGCCCAATGATGCGGATCTTCCAGGAAACGTTGGATTGACAGGA 1073
QY 141 PheIlySerCyArgCyGlyIlyAsnIleuTyrCyValaIleuHisArgTyrSerArgIlyHisThr 160
DB 1074 TTTAAGTGTGGTGGCAACCTTCACTGCGCTTACATCGGTACTCGACAAACACT 1133
QY 161 CysThrTyrArgTyrIlyValaIlyGlnGlnAlaIleAlaIlyAlaAsnProIleuVal 180
DB 1134 TGCACATATGACTACAAAGCCGACGAGGAGAGAGATGCGAAAGCTTATCTTGTTC 1193
QY 181 ValAlaGluIlyValaIlyIlePhe 188
DB 1194 GTGGCCGAGAAAGTTGTCAAGTTT 1217
RESULT 5
ADXS1406 standard; cDNA; 945 BP.
XX ID ADXS1406
XX AC ADXS1406/
DT 21-APR-2005 (first entry)

XX DE Plant full length insert polynucleotide seqid 26146.
XX KW plant protectant; plant growth regulant; gene therapy; pla
XX KW recombinant DNA construct; physical array; plant breeding
XX KW cold tolerance; heat tolerance; drought tolerance; herbic
XX KW extreme osmotic condition; pathogen tolerance; pest tolera
XX KW growth rate; cell cycle pathway; disease resistance; plant
XX KW galactomannan production; lignin production; plant growth
XX KW yield; plant growth; plant development; seed oil; protein
XX KW protein content; gene; ss.
OS Unidentified.
XX PN US2004034888-A1.
XX PD 19-FEB-2004.
XX PD 28-APR-2003; 2003US-00425114.
XX PF 06-MAY-1999; 99US-00304517.
XX PR 05-NOV-2001; 2001US-00985678.
XX PA (LNUJ/) LIU J.
XX PA (ZHOU/) ZHOU Y.
XX PA (KOVA/) KOVALIC D K.
XX PA (SCREEN/) SCREEN S E.
XX PA (TABAS/) TABASKA J E.
XX PA (CAOY/) CAO Y.
XX PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao
XX DR WPI; 2004-180133/17.
XX PT New recombinant DNA construct, useful for improving plant
XX PT cold, heat, drought, herbicides, extreme osmotic condition
XX PT pests, for conferring increased resistance to plant diseases
XX PT improving yield.
XX PS Claim 1, SEQ ID NO 26146; 15pp; English.
XX CC The invention describes a recombinant DNA construct compri
XX CC polynucleotide consisting of a sequence encoding an amino
XX CC available in electronic form from the US patent office at
XX CC ftp://seqdata.nsls.gov/sequence.html?docid:2004034888. The
XX CC of the invention are also useful in physical arrays of mol
XX CC plant breeding markers. The recombinant DNA construct is u
XX CC improving plant tolerance to cold, heat, drought, herbicid
XX CC osmotic conditions, pathogens or pests, for manipulating g
XX CC plant cells by modification of the cell cycle pathway, for
XX CC increased resistance to plant disease, for producing gals
XX CC lignin or plant growth regulators, for increasing the rate
XX CC recombinant in plants, for improving yield by modificat
XX CC photosynthesis or carbohydrate, nitrogen or phosphorus use
XX CC or by providing improved plant growth and development unde
XX CC stress condition or for modifying seed oil or protein yiel
XX CC content. This sequence represents a plant full length inne
XX CC polynucleotide that can be used in the recombinant DNA con
XX CC invention.
SQ Sequence 945 BP; 207 A; 245 C; 256 G; 237 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 3.82e-28 Length: 945
Score: 425.50 Matches: 83
Percent Similarity: 61.67% Conservative: 28
Best Local Similarity: 46.11% Mismatches: 56
Query Match: 43.24% Indels: 13
DB: Gaps: 2
US-10-716-089-20 (1-188) x ADXS1406 (1-945)
QY 8 GlnGluThrThrSerGlnAlaProGlnGlyProValMetCysIlyAsnI

```
Db 112 AAGGAGACTGATGCAGGCACTGAGGACCCATCTTGCATCAATAACTGGGGCTTC 171
Oy 28 PheGlySerGlnAlaThrMetGlyLeuCySerIlySerYrYrArgGluThrValMetGln 47
Db 172 TTTCGCACCGGAGCTACATGAACTGTCCTCCAGTCCCAAGAGATGATGATGAA 231
Oy 48 AlAlaMetThrAlaLeuAlaGluGlnAlaThrGlnAlaAlaGlnAlaThrSerAlaThr 67
Db 232 CAGAGACAGGCGCAAGCTGGCT-----GCTTCCTATTCGACAGATGTCATGGCAAC 285
Oy 68 AlAlaAlaValGlnProProAlaProValHisGluThrIlyLeuThrCysGluValGlu 87
Db 286 GATGCTGTCATGAAACA-----GTTGTTGCT 312
Oy 88 ArgThrMetIleValProHisGlnSerSerSerYrGlnGlnApleuValThrProAla 107
Db 313 GGCAACACAGTGGCTGCTGCTCCATCGAGTTCAACAAATGAACTGACGCCCT 372
Oy 108 AlAlaAlaProGlnAlaValIlySerSerIleAlaAlaProSerArgProGluProAsn 127
Db 373 GATGTTGCTGACCTGACGAGGAGGGGGGGGCTGATCTCCAAAGGAGAGTGGCCCAAC 432
Oy 128 ArgCysGlySerCysArgIlyArgValGlyLeuThrGlyPheIlyCysArgCysGlyAsn 147
Db 433 CGGTGCACACTTTCAGAGAGAGGTTTGACTTACAGATTCAACTGCCGGTGTGGGAAAC 492
Oy 148 LeuTyrCysAlaLeuHisArgTyrSerAspIlyHisIleThrCysThrTyrAspTyrIlyAla 167
Db 493 TTGTACTGTGCACTGCACCGCTACTCCCAAGACGACGACGATGCTGACTATCGGACT 552
Oy 168 AlAlaGlyGlnGluAlaIleAlaIlyValAlaAsnProIleuValValAlaGluIlyValIlyAs 187
Db 553 CTGCTAGAGATGCCATTGCGCAAGCTTAATCCAGTGAAGCGGACAGCAAGTCGACAAAG 612
RESULT 6
AEB66909 standard; DNA; 1000 BP.
AC AEB66909;
XX 22-SEP-2005 (first entry)
XX
DE Rice genome derived DNA sequence, SEQ ID 2054.
XX
XX transcription; gene regulation; transgenic plant; RNA interference;
XX transformation; antibody; ds.
OS Oryza sp.
XX JP2005185101-A.
XX
XX 14-JUL-2005.
XX
XX 11-DEC-2002; 2002JP-00383870.
XX
XX 30-MAY-2002; 2002JP-00203269.
XX
XX (DOKU-) DOKURITSU GYOSEI HOJIN NOGYO SEIBUTSU SH.
XX (SEIB-) SEIBUTSUKI TOKUTEI SANGYO GIJUTSU.
XX (DOKU-) DOKURITSU GYOSEI HOJIN RIKAGAKU KENKYUSH.
XX (KOKU-) ZH KOKUSAI KAGAKU SHIKO ZAIDAN.
XX
XX Kikuchi H, Hayashizaki Y, Otsomo Y, Matsubara K, Murakami K,
XX Kishimoto N, Sato K, Nagata T, Kawakami N, Yasaki J, Ishikawa M;
XX Doi K, Kawai J;
XX
XX WPI, 2005-566181/58.
XX
XX Novel DNA encoding transcription factor, derived from rice plant, useful
XX for obtaining transcriptional-regulatory regions in plant and for
XX producing modified plant.
```

```
PS Claim 1; SEQ ID NO 2054; 2928bp; Japanese.
XX
CC The invention relates to a novel DNA sequence encoding a t:
CC factor derived from a plant. The invention further compris
CC RNA sequences, ribozyme activity RNA, RNAi sequences, a ve
CC transformed plant cells, antibodies and proteins, all relat
CC novel plant DNA sequences of the invention. The novel DNA
CC derived from a rice-genome database. The invention further
CC method for determining the transcription regulatory region;
CC genome. The novel DNA is useful for controlling the expres
CC in a plant and for producing a modified plant with desired
CC characteristics. The plant DNA and method enables the acqu
CC transcriptional-regulatory regions. This polynucleotide rei
CC sequence taken from a rice genome clone library for use in
CC Note: This sequence is not shown in the specification. It i
CC retrieved from a sequence listing in electronic format fro
CC Patent Office. The invention claims DNA sequences of SEQ ID
CC 28469 and encoded protein sequences of SEQ ID Nos 28470 to
CC however, the sequence listing only provided the DNA sequenc
CC Nos 1 to 3032.
XX
SQ Sequence 1000 BP; 234 A; 242 C; 264 G; 260 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 1,38e-27 Length: 1000
Score: 419.50 Matches: 82
Percent Similarity: 60.00% Conservative: 26
Best Local Similarity: 45.56% Mismatches: 59
Query Match: 42.63% Indels: 13
DB: 14 Gaps: 2
US-10-716-089-20 (1-188) x AEB66909 (1-1000)
Oy 8 GlnGluThrThrSerGlnAlaProGluGlyProValMetCysIlyAsnI 27
Db 160 AAGGAGACTGATGCCAGACCCAGAGGCCCGGATCTTGCATCAATA 219
Oy 28 PheGlySerGlnAlaThrMetGlyLeuCySerIlySerYrYrArgGlu 47
Db 220 TTTCGCACCGGAGCTACATGAACTGTCCTCCAGTCCCAAGAGAC 279
Oy 48 AlAlaMetThrAlaLeuAlaGluGlnAlaThrGlnAlaAlaGlnAla 67
Db 280 CAGAGACAGGCGCAAGCTGGCA-----GCTTCCTATTCGACAGATG 333
Oy 68 AlAlaAlaValGlnProProAlaProValHisGluThrIlyLeuThrC 87
Db 334 GATTCGCGAAGGAAACCAATTATGCTGTCACCGCTGAAGTACGTTC 393
Oy 88 ArgThrMetIleValProHisGlnSerSerSerYrGlnGlnApleuV 107
Db 394 GTGAAGACGCTTGT-----G 420
Oy 108 AlAlaAlaProGlnAlaValIlySerSerIleAlaAlaProSerArgP 127
Db 421 GAGATTGCTGCTCTATGAGGGGTCACGCTGAACCCCAAGGAGAGG 480
Oy 128 ArgCysGlySerCysArgIlyArgValGlyLeuThrGlyPheIlyCys 147
Db 481 CGGTGCACACTTTCAGAGAGAGGTTTGACTTACAGATTCAACTGCC 540
Oy 148 LeuTyrCysAlaLeuHisArgTyrSerAspIlyHisIleThrCys 167
Db 541 TTGTACTGTGCACTGCACCGCTACTCCAGATGACATGACATGCTG 600
Oy 168 AlAlaGlyGlnGluAlaIleAlaIlyValAlaAsnProIleuVal 187
Db 601 CTGCTAGAGATGCTATTCGCAAGCTTAATCCGCTGCTGGAAGCGGAGA 660
RESULT 7
AEB66932 standard; DNA; 1021 BP.
XX
```


AC AEB66932;
XX 22-SEP-2005 (first entry)
XX
XX Rice genome derived DNA sequence, SEQ ID 2077.
DE
XX transcripction, gene regulation; transgenic plant; RNA interference;
XX transformation; antibody, ds.
XX
OS Oryza sp.
XX
XX JP2005185101-A.
PN
XX 14-JUL-2005.
PD
XX 11-DEC-2002; 2002JP-00383870.
PP
XX 30-MAY-2002; 2002JP-00203269.
PR
XX (DOKU-) DOKURITSU GYOSEI HOJIN NOGYO SEIBUTSU SH.
PA (SEIB-) SEIBUTSUKEN TOKUKEI SANJO GIUTSU.
PA (DOKU-) DOKURITSU GYOSEI HOJIN RIKKAGAKU KENKUSH.
XX (KOKU-) ZH KOKUSAI KAGAKU SHINKO ZAIDAN.
XX
PI Kikuchi H, Hayashizaki Y, Ootomo Y, Matsubara K, Murakami K;
PI Kishimoto N, Sato K, Negata T, Kawakami N, Yazaki J, Ishikawa M;
PI Doi K, Kawai J,
DR
XX WPI, 2005-566181/58.
PT Novel DNA encoding transcription factor, derived from rice plant, useful
PT for obtaining transcriptional-regulatory regions in plant and for
PT producing modified plant.
XX
XX Claim 1, SEQ ID NO 2077; 2928bp; Japanese.
PS
XX The invention relates to a novel DNA sequence encoding a transcription
CC factor derived from a plant. The invention further comprises antisense
CC RNA sequences, ribozyme activity RNA, RNAi sequences, a vector,
CC transformed plant cells, antibodies and proteins, all related to the
CC novel plant DNA sequences of the invention. The novel DNA is preferably
CC derived from a rice-genome database. The invention further provides a
CC method for determining the transcription regulatory regions of the rice
CC genome. The novel DNA is useful for controlling the expression of a gene
CC in a plant and for producing a modified plant with desired and different
CC characteristics. The plant DNA and method enables the acquisition of many
CC transcriptional-regulatory regions. This polynucleotide represents a DNA
CC sequence taken from a rice genome clone library for use in the invention.
CC Note: This sequence is not shown in the specification. It has been
CC retrieved from a sequence listing in electronic format from the Japanese
CC Patent Office. The invention claims DNA sequences of SEQ ID Nos 1 to
CC 28469 and encoded protein sequences of SEQ ID Nos 28470 to 56791,
CC however, the sequence listing only provided the DNA sequences of SEQ ID
CC Nos 1 to 3032.
XX
XX Sequence 1021 BP; 240 A; 244 C; 269 G; 268 T; 0 U; 0 Other;
XX
XX
XX Alignment Scores:
Pred. No.: 1,42e-27 Length: 1021
Score: 419.50 Matches: 82
Percent Similarity: 60.00% Conservative: 26
Best Local Similarity: 45.56% Mismatches: 59
Query Match: 42.63% Indels: 13
DB: 14 Gaps: 2
US-10-716-089-20 (1-188) x AEB66932 (1-1021)
Qy 8 GlnGluThrThrsSerGlnAlaProGluGlyProValMetCysGlyAsnLeuCysGlyPhe 27
Dy 156 AAGGAGACGACGAGCCAGCAGCCAGAGAGCCGAGATCCTTTCGATCAATAAATCGCGCTTC 215
Yy 28 PheGlySerGlnAlaThrMetCysGlySerGlyCysTyrArgGluThrValMetCln 47

Db 216 TTTCGAGTCGGCGCTACCATGAACATGTGCTCAAAATGGCCACAAAGAGA
Qy 48 AATATyMeTThrAlaLeuAlaGlnAlaThrGlnAlaAlaGlnAlaT
Db 276 CAGGAGCAGGCCAAGCTGGCA-----GCCCTCCTTATCCAGACATTTG
Qy 68 AAlaAlaAlaValGlnProProAlaProValHisGlnThrLysLeuThrC
Db 330 GATTCGCGGAGAGAAACCAATTATTGCTGTGCACGCTGAAGTACGTGTG
Qy 88 ArgThrMeIleValProHisGlnSerSerSerTyrGlnGlnAspLeuV
Db 390 GTGAAGACCGCTGTT-----G
Qy 108 AAlaAlaAlaProGlnAlaValLysSerSerIleAlaAlaProSerArgP
Db 417 GAGATTGCTGTCCTTACTGTAGGGGGTCCAGCGTAAACCCCAAGGAGGG
Qy 128 ArgCysGlySerCysArgLysArgValGlyLeuThrGlyPheLysCysA
Db 477 CGGTGCTCCACTGTTCGGAAGAGGGGTGGCTTAACCGATTCAACTGCC
Qy 148 LeuTyrCysAlaLeuHisArgTyrTyrSerAspLysHisThrCysThrTyrA
Db 537 TTGTACTGTGGCAATGCACCGCTATTCCGATTAACATGACATGACCGCAGTTG
Qy 168 AAlaGlnGlnAlaAlaIleAlaLysAlaAsnProLeuValValAlaGlnL
Db 597 GCTGCTAAGGATGCTATTGTCCAAGGCTAATCCGCTGTGTGAAGCGCGAGA
RESULT 8
ACLS0227
ID ACLS0227 standard; cDNA; 519 BP.
XX ACLS0227;
XX AC
XX DT 02-JUN-2005 (first entry)
XX XX
XX DE Rice abiotic stress responsive polynucleotide SEQ ID NO:83
XX KM 89; abiotic stress tolerance; transgenic plant; plant; cer
XX KM agriculture.
XX OS Oryza sativa.
XX XX
XX PN WO2003008540-A2.
XX PD 30-JAN-2003.
XX PF 21-JUN-2002; 2002WO-US019668.
XX XX
XX PR 22-JUN-2001; 2001US-0300112P.
XX PR 24-AUG-2001; 2001US-0314662P.
XX PR 26-SEP-2001; 2001US-0325277P.
XX PR 21-NOV-2001; 2001US-0332312P.
XX PA (SYGN) SYNGENTA PARTICIPATIONS AG.
XX XX
XX P1 Krieps J, Briggs SP, Cooper B, Glazebrook J, Goff SA,
XX P1 Moughamer T, Provart N, Ricke D, Zhu T,
XX XX
XX DR WPI; 2003-248011/24.
XX XX
XX PT New stress-responsive nucleic acid, useful for altering th
XX PT responsiveness of a plant, e.g. cereal, to an abiotic stre
XX PT stress; salt stress or osmotic stress.
XX XX
XX PS Claim 1; SEQ ID NO 8314; 89pp; English.
XX XX

could
less
part

jes
cat

CC responsiveness of a plant to abiotic stress. The invention is useful in
 CC agriculture. The nucleic acid is useful for determining whether a test
 CC plant has been exposed to an abiotic stress condition. It is also useful
 CC for selecting an agent that alters abiotic stress regulated
 CC polynucleotide expression in a plant cell, and to identify a homolog or
 CC ortholog to an abiotic stress responsive polynucleotide. The nucleic acid
 CC molecule and the polypeptide encoded by it are useful in altering the
 CC responsiveness of a plant to an abiotic stress, such as cold stress, salt
 CC stress, osmotic stress or any of their combinations. The present sequence
 CC is used in the exemplification of the invention

XX Sequence 519 BP; 117 A; 133 C; 172 G; 97 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1,14e-27	Length:	519
Score:	416.50	Matches:	82
Percent Similarity:	57.29%	Conservative:	28
Best Local Similarity:	42.71%	Mismatches:	47
Query Match:	42.33%	Indels:	35
DB:	11	Gaps:	3

US-10-716-089-20 (1-188) x ACL30227 (1-519)

```

QY      8 GlnGluThrThrSerGlnAlaProGluGlyProValMetCysIysAsnLeuCysGlyPhe 27
      10 AAGGAGGCTGGCTGCAGAGCCGAGGGCCCAATCTTTGCATCAATTAATGCGCGCTTC 69
DB      28 PheGlySerGlnAlaThrMetGlyLeuCysSerIysCysYrYrArgGluThrValMetGln 47
      70 TTTGGCAGCGCGCGGCACCATGAACTGTGCTCCAAAGTCCCAAGAGATGATCATGTGAAG 129
QY      48 AlAluMetThrAlaLeuAlaGluGlnAlaThrGlnAla----- 60
      130 GAGGAGCAGGCCAGACTTCTCCTCCATCATGATGATGATGATGATGATGATGATGATGAT 189
DB      61 -----AlaGlnAlaThrSerAlaThrAlaAlaAlaAlaAlaValGlnProProAla 75
      190 GGGAGAGCATATTGTTGCTGCAGTGTGACAGCGCGTGGCGGTG----- 237
QY      76 ProValHisGluThrIleuThrCysGluValGluArgThrMetIleValProHisGln 95
      238 -----GCTCAGGTGAGGGGAGAGACCGCTCGTT----- 264
DB      96 SerSerSerTyrglnGlnAspLeuValThrProAlaAlaAlaAlaProGlnAlaValIys 115
      265 -----GTGCAAGCTTACCGATGTGCGGGGACACGACGAGAG 300
QY      116 SerSerIleAlaAlaProSerArgProGluProAsnArgCysGlySerCysArgIysArg 135
      301 GTTGTCTGTAAGTCCCAAGGTCAAGAGGCGGACCCGATGCTGTAAGAGAG 360
DB      136 ValGlyLeuThrGlyPheIysCysArgCysGlyAsnLeuTyrcyAlaLeuHisArgIyr 155
      361 GTTGGGCTGACCGGATTCACATGCGCGGTGCGGTAACTGTACTGTGCGGTGACCGCTAC 420
QY      156 SerAspIysHisLeuThrCysThrTyraAspTyrlsAsaIleAlaGlyGlnGluAlaIleAlaIys 175
      421 TCCGCAAGCATGAATGCCAGTTCACTACCGAGACTGCGGCTAGGAGATGCCATGCCAAG 480
DB      176 AlaAsnProLeuValAlaAlaGluIysValIys 187
      481 GCCAACCACGATGCTGAAGCTGAGAAAGCTGCACAG 516

```

RESULT 9

ID ADA70637 standard; DNA; 522 BP.

AC ADA70637;

XX 20-NOV-2003 (first entry)

DT Rice gene, SEQ ID 3960.

DE

XX

KW Plant; bacterial infection; fungal infection; viral infecti
 KW gene; ds.

OS Oryza sativa.

PN WO2003000898-A1.

PD 03-JAN-2003.

XX 22-JUN-2001; 2001WO-IB001105.

XX 22-JUN-2001; 2001WO-IB001105.

XX 22-JUN-2001; 2001WO-IB001105.

XX (SYGN) SYNGENTA PARTICIPATIONS AG.

XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou

XX F, Katagiri F, Qian S, Tao Y, Whitham S, Xie Z, Zhu T, ;

XX WPI; 2003-175290/17.

XX WPI; 2003-175290/17.

XX Claim 6; SEQ ID NO 3960; 899pp; English.

XX The present invention relates to a method (M1) for identifi
 CC involved in plant resistance or response to pathogenic inf
 CC comprises identifying a gene whose expression is significan
 CC the incompatible interaction of plant gene expression relat
 CC the expression of the gene in an uninfected plant, in a mutant
 CC or in a corresponding incompatible or compatible interacti
 CC useful for conferring resistance to resistance or toleranc
 CC bacterial, fungal or viral infection. The present sequence
 CC illustrate the invention.

XX sequence 522 BP; 119 A; 133 C; 172 G; 98 T; 0 U; 0 Other;

XX Alignment Scores:

Pred. No.:	1,14e-27	Length:	522
Score:	416.50	Matches:	82
Percent Similarity:	57.29%	Conservative:	28
Best Local Similarity:	42.71%	Mismatches:	47
Query Match:	42.33%	Indels:	35
DB:	8	Gaps:	3

US-10-716-089-20 (1-188) x ADA70637 (1-522)

```

QY      8 GlnGluThrThrSerGlnAlaProGluGlyProValMetCysIysAsnL 27
      10 AAGGAGGCTGGCTGCAGAGCCGAGGGCCCAATCTTTGCATCAATTAATGCGCGCTTC 69
DB      28 PheGlySerGlnAlaThrMetGlyLeuCysSerIysCysYrYrArgGluTr 47
      70 TTTGGCAGCGCGCGGCACCATGAACTGTGCTCCAAAGTCCCAAGAGATG 129
QY      48 AlAluMetThrAlaLeuAlaGluGlnAlaThrGlnAla----- 60
      130 GAGGAGCAGGCCAAGCTTCTCCTCCATCATGATGATGATGATGATGATGATGATGATGAT 189
DB      61 -----AlaGlnAlaThrSerAlaThrAlaAlaAlaAlaValG 75
      190 GGGAGAGCATATTGTTGCTGCAGTGTGACAGCGCGTGGCGGTG----- 237
QY      76 ProValHisGluThrIleuThrCysGluValGluArgThrMetIleV 95
      238 -----GCTCAGGTCCAGGCGGAGAGACGCTGC 264
DB      96 SerSerSerTyrglnGlnAspLeuValThrProAlaAlaAlaAlaProG 115
      265 -----GTGCAAGCTTACCGATGTGCGGGGACACGACGAGAG 300

```

```
QY 116 SerSerIleAlaIaIaProSerArgProGluProAenArgCysGlySerCysArgIysArg 135
Db 301 GTTGGCTGTAGTCCCAAGGTCAGAGAGGCCCAACCGGTCGCTTCCTGTGAGAGAG 360
QY 136 ValGlyLeuThrGlyPheIysCysArgCysGlyAsnLeuTyrCysAlaLeuHisArgTyr 155
Db 361 GTTGGGCTGACGGGATTCACACTCCCGTCGGTAAACATGATCTGTCCTTGCACCGCTAC 420
QY 156 SerAspIysHisIleThrCysThrTyrAspTyrLysAlaIaGlyGlnGlnIaIaIaIaIys 175
Db 421 TCCGACAAACATGAATGACCGGTCGACTACCGGACGCGGCTAGGATGCCATCCGCAAG 480
QY 176 AlaAsnProLeuValValAlaGluLysValIys 187
Db 481 GCCAACCCAGTGTGAGGCTGAGAGCTCGACAAAG 516

RESULT 10
ACL27951
ID ACL27951 standard; cDNA; 522 BP.
AC ACL27951;
XX
XX
XX 02-JUN-2005 (first entry)
DT
XX
XX Rice abiotic stress responsive polynucleotide SEQ ID NO:1907.
XX
XX ss; abiotic stress tolerance; transgenic plant; plant; cereal;
XX agriculture.
XX
XX Oryza sativa.
XX
XX WO2003008540-A2.
XX
XX 30-JAN-2003.
XX
XX 21-JUN-2002; 2002WO-US019668.
XX
XX 22-JUN-2001; 2001US-0300112P.
XX 24-AUG-2001; 2001US-0314662P.
XX 26-SEP-2001; 2001US-0325277P.
XX 21-NOV-2001; 2001US-0332132P.
XX
XX (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX
XX Kreps J, Briggs SP, Cooper B, Glazebrook J, Goff SA, Katagiri F,
XX Moughamer T, Provart N, Rieke D, Zhu T,
XX
XX WPI; 2003-248011/24.
XX
XX New stress-responsive nucleic acid, useful for altering the
XX responsiveness of a plant, e.g. cereal, to an abiotic stress such as cold
XX stress, salt stress or osmotic stress.
XX
XX Claim 1, SEQ ID NO 1907; 89pp; English.
XX
XX The invention relates to novel abiotic stress responsive polynucleotides
XX and polypeptides. Also disclosed are vectors, expression cassettes, host
XX cells, and plants containing such polynucleotides. Also disclosed are
XX methods for using the polynucleotides and polypeptides to alter the
XX responsiveness of a plant to abiotic stress. The invention is useful in
XX agriculture. The nucleic acid is useful for determining whether a test
XX plant has been exposed to an abiotic stress condition. It is also useful
XX for selecting an agent that alters abiotic stress regulated
XX polynucleotide expression in a plant cell, and to identify a homolog or
XX ortholog to an abiotic stress responsive polynucleotide. The nucleic acid
XX molecule and the polypeptide encoded by it are useful in altering the
XX responsiveness of a plant to an abiotic stress, such as cold stress, salt
XX stress, osmotic stress or any of their combinations. The present sequence
XX is used in the exemplification of the invention
XX
XX Sequence 522 BP; 119 A; 133 C; 172 G; 98 T; 0 U; 0 Other;
XX
XX
```

```
Alignment Scores:
Pred. No.: 1,14e-27 Length: 522
Score: 416.50 Matches: 82
Percent Similarity: 57.29% Conservative: 28
Best Local Similarity: 42.71% Mismatches: 47
Query Match: 42.33% Indels: 35
DB: 11 Gaps: 3

US-10-716-089-20 (1-188) x ACL27951 (1-522)
QY 8 GlnGluThrThrSerGlnAlaProGluGlyProValMetCysLeuAsnI 27
Db 10 AAGGAGGCTGTGCTGCCAGCAGCCGACGCCCAATCCTTGTGATCAAT 69
QY 28 PheGlySerGlnAlaThrMetGlyLeuCysSerIysCysTyrArgGlu 47
Db 70 TTTGGCAGCGCGCGACCATGAAATGTCCTCAAGTCGACCAAGAGAG 129
QY 48 AlaLysMetThrAlaLeuAlaGluGlnIaIaThrGlnAla----- 60
Db 130 GAGGAGCAGCCCAAGCTTGTGCTCTCTCCATCATGATGATTCATG 189
QY 61 -----AlaGlnAlaThrSerAlaThrAlaAlaValG 75
Db 190 GGAAGAGACATATTGTTGCTGCCAGTGTACACGCGCGCGCTG- 237
QY 76 ProValHisGluThrLysLeuThrCysGluValGluArgThrMetIle 95
Db 238 -----GCTCAGGTGCAAGGCAAGACGCTG 264
QY 96 SerSerSerTyrGlnGlnAspLeuValThrProAlaAlaAlaIaPro 115
Db 265 -----GTGCAGCTTACCGATGTGCGCGGCA 300
QY 116 SerSerIleAlaIaProSerArgProGluProAenArgCysGlySer 135
Db 301 GTTGGCTGTAGTCCCAAGGTCAGAGAGGCCCAACCGGTCGCTACCT 360
QY 136 ValGlyLeuThrGlyPheIysCysArgCysGlyAsnLeuTyrCysAla 155
Db 361 GTTGGGCTGACGGGATTCACACTCCCGTCGGTAAACATGATCTGTC 420
QY 156 SerAspIysHisIleThrCysThrTyrAspTyrLysAlaIaGlyGln 175
Db 421 TCCGACAAACATGAATGACCGGTCGACTACCGGACGCGGCTAGGAT 480
QY 176 AlaAsnProLeuValValAlaGluLysValIys 187
Db 481 GCCAACCCAGTGTGAGGCTGAGAGCTCGACAAAG 516

RESULT 11
ADXS2834
ID ADXS2834 standard; cDNA; 1032 BP.
XX
XX ADXS2834;
XX
XX 21-APR-2005 (first entry)
XX
XX Plant full length insert polynucleotide seqid 27574.
XX
XX plant protectant; plant growth regulant; gene therapy; pla
XX recombinant DNA construct; physical array; plant breeding
XX cold tolerance; heat tolerance; drought tolerance; herbic
XX extreme osmotic condition; pathogen tolerance; pest tolera
XX growth rate; cell cycle pathway; disease resistance;
XX galactomannan production; lignin production; plant growth
XX yield; plant growth; plant development; seed oil; protein
XX protein content; gene; ss.
XX
XX unidentified.
XX
XX US2004034888-A1.
XX
XX
```


CC responsiveness of a plant to an abiotic stress, such as cold stress, salt
CC stress, osmotic stress or any of their combinations. The present sequence
CC is used in the exemplification of the invention

Sequence 522 BP; 122 A; 144 C; 168 G; 88 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.55e-27	Length:	52
Score:	415.00	Matches:	85
Percent Similarity:	61.88%	Conservative:	27
Best local Similarity:	46.96%	Mismatches:	57
Query Match:	42.17%	Indels:	12
GB:	11	Gaps:	4

US-10-716-089-20 (1-188) x ACL34078 (1-522)

Oy		GlnGluThrThrsSerGlnAlaProGluGlyProValMetCysAluSeraLeuLysCySeGlyIype	27
Db			
Oy		10 AAGGAAGCTGGGTGCACGAGCCCGAAGGAAACCATCTCTCATCAATAACTGTGGCTTC	69
Db			
Oy		28 PheGlySerGlnAlaThrMetGlyLeuCySeSerLysCysTyraArgJuthrValMetGln	47
Db			
Oy		70 TTCGGCAGCGCGCGCACCATGAACATTGTGCTCCAAAGTGCACAAAGGAGATGATTAAGAAG	129
Db			
Oy		48 AlaLysMetThrAlaLeuAlaGluGlnAlaThrGlnAlaAlaGlnAlaThrSeraAlaTrn	67
Db			
Oy		130 CAGGATCAGGCCAACGTGGCT-----GCCTCTCTTATGACACACATCGTAACGCGCAGC	183
Db			
Oy		68 AlaAlaAlaValaGlnProProAlaProValHisGluThrLysLeuThrCysGluValaGlu	87
Db			
Oy		184 GACCGCGTCATGAGACCGCGTTTGCTGCGACGAACAAG-----GTA GTA	228
Db			
Oy		88 ArgThrMetIleValProHisGlnSerSerSerTyrgInaIlePLeuValTrnProAla	107
Db			
Oy		229 GCGTGTGCCCAAGTTGAGTTGCAAAACATGAACGTCGACGAG-----CCCGCT	276
Db			
Oy		108 ---AlaAlaAlaProGlnAlaValaLysSerSerIleAlaAlaProSeraArgProGluPro	126
Db			
Oy		277 GATGTTGCCGAGACCCGACGAGGGGGGTGCGCGGAGATCTCAAAGGGGGGAAGTAGGGCGG	336
Db			
Oy		127 AsnArgCysGlySerCysArgLysArgValaGlyLeuThrGlyPheLysCysArgCysGly	146
Db			
Oy		337 AACCGGTGCAGCGCCCTGCAGAAAGAGGGTGGACTTACGGGATTCACCTGCCGTGTGGG	396
Db			
Oy		147 AsnLeuTyrcysAlaLeuHisArgTyriserAspLysHisIleThrCysThrTyraSPtyrLys	166
Db			
Oy		397 AACTGTACTGCGCAGCTCCACCGCTACTCCGACAAGCAGCAGCTCAAGATTGCAGTACC	456
Db			
Oy		167 AlaAlaGlyGlnGlnAlaIleAlaValaAlaAspValaIleAlaGluLysValaVala	186
Db			
Oy		457 ACTGCTGCACAGGAGCCCATTTGCCAAGGCTAATTCGGGTGAAGAGCAACAACTCGAC	516
Db			
Oy		187 Lys	187
Db		517 AAG	519

RESULT 13

ID ADX47307 standard; cDNA; 736 BP.

AC ADX47307,

XX 21-APR-2005 (first entry)

DE Plant full length insert polynucleotide seqid 22047.

KM plant protectant; plant growth regulant; gene therapy; plant;

KM recombinant DNA construct; physical array; plant breeding marker;

KM cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;

KW extreme osmotic condition; pathogen tolerance; pest tolerance;

KW growth rate; cell cycle pathway; disease resistance;

KW galactonman production; lignin production; plant growth regulator;

KW yield; plant growth; plant development; seed oil; protein yield;

KW protein content; gene; ss.

OS Unidentified.

PN US2004034888-A1.

PD 19-FEB-2004.

PF 28-APR-2003; 2003US-00425114.

06-MAY-1999: 99US-00304517.

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

PA (ZHOU/) ZHOU Y

PA (NOVA/) SCREEN S.E.

PA	(1ABH/)	1ABASHA U E.
PA	(CAOY/)	CAO Y

XX	Tju J	Zhou Y	Kou
DT			

XX
WDT: 2004 100123/17

XX **Now, roughly, what data**

PT cold, heat, drought, herbicides, extreme osmotic conditions

PT improving yield.

PS Claim 1; SEQ ID NO 22047; 15pp; English.

CC The invention describes a recombinant DNA construct comprising

CC available in electronic form from the US patent office at

of the invention

improving plant tolerance to cold, heat,

CC osmotic conditions, pathogens or pests, for manipulating
CC plant cells by modification of the cell cycle pathway, for
CC increased resistance to plant disease, for producing galactose
CC lignin or plant growth regulators, for increasing the rate
CC recombination in plants, for improving yield by modifica-
CC photosynthesis or carbohydrate, nitrogen or phosphorus use
CC or by providing improved plant growth and development unde
CC stress condition or for modifying seed oil or protein yield
CC content. This sequence represents a plant full length insere
CC polynucleotide that can be used in the recombinant DNA con
CC invention.

Sequence 736 BP; 158 A; 213 C; 225 G; 140 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2,376-27	Length:	726
Score:	415.00	Matches:	85
Percent Similarity:	61.88%	Conservative:	27
Best Local Similarity:	46.96%	Mismatches:	57
Query Match:	42.17%	Indels:	12
AB:	13	Gaps:	4

US-10-716-089-20 (1-188) x ADX47307 (1-736)

QY	8	GInGluThrThrSerGlnAlaProGluGlyProValMetCysLysAsnL	27
	113	AAGAGAGCTGGGTGCGAGGCCCGAGGAGGCCATCCTCTGCATCAATTA	172
Db	28	PheGlySerGlnAlaThrMetGlyLeuCysSerIleCysIleArgGluT	47
	173	TTCGGCAGCGCGGACCATGAACATGTGCTCCAAGTGCCACAAGAGAGA	232
QY	48	AlaLysMetThrAlaLeuAlaGluGlnAlaThrGlnAlaAlaGlnAlaT	67
	233	CAGATCAGAGCCCAAGCTGGCT-----GCTTCCTCTATGCAAGCATCG	286
Db			

DB 640 AAG 642
RESULT 15
AD47342
ID AD47342 standard; cDNA; 821 BP.
XX
XX AD47342;
DT 21-APR-2005 (first entry)
XX
XX plant full length insert polynucleotide seqid 22082.
XX
XX plant protectant; plant growth regulant; gene therapy; plant;
XX recombinant DNA construct; physical array; plant breeding marker;
XX cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
XX extreme osmotic condition; pathogen tolerance; pest resistance;
XX growth rate; cell cycle pathway; disease resistance;
XX galactomannan production; lignin production; plant growth regulator;
XX yield; plant growth; plant development; seed oil; protein yield;
XX protein content; gene; 88.
OS Unidentified.
XX
XX US2004034888-A1.
XX
XX 19-FEB-2004.
XX
XX 28-APR-2003; 2003US-00425114.
XX
XX 06-MAY-1999; 99US-00304517.
XX
XX 05-NOV-2001; 2001US-00985678.
XX
XX (LIUJ) LIU J.
XX (ZHOU) ZHOU Y.
XX (KOVA) KOVALIC D K.
XX (SCRE) SCREEN S E.
XX (TABA) TABASKA J E.
XX (CAOY) CAO Y.
XX
XX Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
XX WPI; 2004-180133/17.
XX
XX New recombinant DNA construct, useful for improving plant tolerance to
XX cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
XX pests, for conferring increased resistance to plant disease, or for
XX improving yield.
XX
XX Claim 1, SEQ ID NO 22082; 15pp; English.
XX
XX The invention describes a recombinant DNA construct comprising a
XX polynucleotide consisting of a sequence encoding an amino acid sequence
XX available in electronic form from the US patent office at
XX ftp://seqdata.uspto.gov/sequence.html?docid:2004034888. The polynucleotide
XX of the invention are also useful in physical arrays of molecules and as
XX plant breeding markers. The recombinant DNA construct is useful for
XX improving plant tolerance to cold, heat, drought, herbicides, extreme
XX osmotic conditions, pathogens or pests, for manipulating growth rate in
XX plant cells by modification of the cell cycle pathway, for conferring
XX increased resistance to plant disease, for producing galactomannan,
XX lignin or plant growth regulators, for increasing the rate of homologous
XX recombination in plants, for improving yield by modification of
XX photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
XX or by providing improved plant growth and development under at least one
XX stress condition or for modifying seed oil or protein yield and/or
XX content. This sequence represents a plant full length insert
XX polynucleotide that can be used in the recombinant DNA construct of the
XX invention.
XX
XX Sequence 821 BP, 171 A, 236 C, 247 G, 167 T, 0 U, 0 Other;
XX
Alignment Scores: 2.71e-27 Length: 821
Pred. No.: 2.71e-27 Length: 821

Score: 415.00 Matches: 85
Percent Similarity: 61.88% Conservative: 27
Best Local Similarity: 46.96% Mismatches: 57
Query Match: 42.17% Indels: 12
DB: 13 Gaps: 4
US-10-716-089-20 (1-188) x AD47342 (1-821)
QY 8 GlnGlnThrThrSerGlnAlaProGlnGlyProValMetCysValAsnL 27
DB 113 AAGAGGCTGGGTGCGGAGGCGCCCGAGGAGCCCATCTGTCATATTA 172
QY 28 PheGlySerGlnAlaThrMetGlyLeuCysSerLeuSerCysTyrArgGluT 47
DB 173 TTCGGCAGCGCGCGCCACATGAACATGTGTCCCAAGTCCCAAGAGAG 232
QY 48 AlaValMetThrAlaLeuAlaGlnGlnAlaThrGlnAlaAlaGlnAlaT 67
DB 233 CAGGATCAGGCCCAAGCTGGT-----GCTTCCTATATCAGACATTC 286
QY 68 AlaAlaAlaValGlnProProAlaProValHisGlnThrLeuThrC 87
DB 287 GACGCCGTCAATGAGCCGCTGTGCTGTCGACAGCAACACG----- 331
QY 88 ArgThrMetIleValProHisGlnSerSerSerTyrGlnGlnAspLeu 107
DB 332 GCTGTGCCCAAGTTGAAGTTGCACAAATGAACGTGCAGCAG----- 379
QY 108 --AlaAlaAlaProGlnAlaValIleSerSerSerIleAlaAlaProSer 126
DB 380 GATGTGCCGAGCCAGCCAGGAGGGGTGGCGGATCTCCAAAGGGGGA 439
QY 127 AsnArgCysGlySerCysArgLeuArgValGlyLeuThrGlyPheLeu 146
DB 440 AACCGGTCCAGCGCCTGCAGAAAGAGGGTTGAGATTACGGATTCACG 499
QY 147 AsnLeuTyrCysAlaLeuHisArgTyrSerArgLeuValIsthrCysThrT 166
DB 500 AACTGTACTGCGCACTCCACCGCTACTCCGACCAAGCAGCACTGCAAGT 559
QY 167 AlaAlaGlyGlnGlnAlaIleAlaValAsnProLeuValAlaIAG 186
DB 560 ACTGCTGCCAGAGGAGCCATTGCCAAGGCTAATCCGTGTGAAGGCAG 619
QY 187 Lys 187
DB 620 AAG 622

Search completed: December 8, 2005, 16:44:37
Job time : 477 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 8, 2005, 16:33:58 ; Search time 3752 Seconds
(without alignment)
2344.342 Million cell updates/sec

Title: US-10-716-089-20

Perfect score: 984
Sequence: 1 MATERVSOETTSQAEQPV.....GGEALAKNPVLAKEVKVF 188

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame_p2n.model -DEV=xlp
-Q=/cgn2.1/USPTO.spool.p/US10716089/runat.05122005.094344.13151/app.query.fasta_1.327
-DB=EST -OPMT=fastlap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPTOL=0 -LOOPEXT=0
-UNITS=bits -STRAT=1 -NDS=1 -MATRIX=blosum62 -TRANS=human4.cdi -LIST=45
-DOCAI=200 -THR_SCORE=PCT -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-USER=US10716089@cgn2.1.8010@runat.05122005.094344.13151 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEORDER -NEG_SCORES=0 -WAIT -DSELOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -NARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST.*
1: gb_est1.*
2: gb_est2.*
3: gb_est3.*
4: gb_est4.*
5: gb_est5.*
6: gb_est6.*
7: gb_est7.*
8: gb_est8.*
9: gb_est9.*
10: gb_est10.*
11: gb_est11.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	932	94.7	755	3	BJ958521 BJ958521
C 2	840	85.4	769	3	BJ610653 BJ610653
C 3	834	84.8	729	3	BJ592532 BJ592532
C 4	834	84.8	733	3	BJ592791 BJ592791
C 5	825	83.8	747	3	BJ584946 BJ584946
C 6	824	83.7	727	3	BJ585055 BJ585055
C 7	804	81.7	721	3	BJ585014 BJ585014

C 8	786	79.9	702	3	BJ606134
C 9	777	79.0	704	3	BJ591761
C 10	708	72.0	719	3	BJ941620
C 11	687	69.8	664	3	BJ165418
C 12	682	69.3	738	3	BJ605573
C 13	676	68.7	723	3	BJ598461
C 14	668	67.9	736	3	BJ586048
C 15	663	65.3	723	3	BJ611574
C 16	617	62.7	642	3	BJ602584
C 17	602	61.2	672	3	BJ961154
C 18	573	58.2	670	3	BJ590372
C 19	564	57.3	651	3	BJ961357
C 20	564	57.3	652	3	BJ952449
C 21	560	56.9	589	3	BJ173073
C 22	546	55.5	612	5	BJ052287
C 23	529	53.8	644	3	BJ601969
C 24	524	53.3	644	3	BJ597686
C 25	524	53.3	644	3	BJ599109
C 26	516	52.4	651	3	BJ602208
C 27	472	48.0	606	3	BJ167125
C 28	451	45.8	760	7	CN208865
C 29	449	45.6	932	7	CV735162
C 30	448	45.5	708	3	BJ950422
C 31	446	45.3	726	3	BJ950225
C 32	438.5	44.6	534	7	CK757250
C 33	436.5	44.4	716	8	DR102626
C 34	436.5	44.4	740	8	DR387016
C 35	436.5	44.4	743	8	DR025280
C 36	436.5	44.4	746	6	CF401326
C 37	436.5	44.4	798	6	CF395789
C 38	436.5	44.4	802	6	CF671829
C 39	436.5	44.4	809	6	DR048233
C 40	436.5	44.4	818	6	CF402491
C 41	436.5	44.4	820	7	CO165465
C 42	436.5	44.4	837	8	DR101171
C 43	436.5	44.4	865	8	DR014994
C 44	436.5	44.4	884	8	DR052935
C 45	436.5	44.4	892	8	DR014741

ALIGNMENTS

RESULT 1
BJ958521/c 755 bp mRNA linear
LOCUS BJ958521 pphf full-length cDNA library Physcomitrel...
DEFINITION patens cDNA clone pphf6m05 3', mRNA sequence.
ACCESSION BJ958521
KEYWORDS BJ958521.1 GI:67698288
SOURCE EST.
ORGANISM Physcomitrella patens subsp. patens
Bryophyta: Vriidiplantae; Streptophyta; Embryophyta
Bryopsida; Funariidae; Funariales; Funariaceae; ph
1 (bases 1 to 755)
Fujita,T., Nishiyama,T., Shin-i,T., Kohara,Y. and H
Physcomitrella patens EST at a stage of the first a
division of protoplasts
Unpublished (2005)
Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yatae, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshin@genes.nig.ac.jp

Protoplasts were inoculated on BCDMG medium for ev
Protoplasts were isolated from the protomata, fur
at 25C under continuous light for 2-3 days. The reg
which were rich in cells at a stage during the first
cell division, were collected. Total RNA was extrac
constructing a full-length cDNA library. The datab
clones is available at the PHYSCDBase (http://mo88.


```

Db      648 GCTGCGCTGTTACGCCCCCGCTCTGTCATGACACCAAGTCACATCGAGGTTGAG 589
Qy      88 ArgThmCtIleVal1ProH1gInSerSerSerTyTgInGlnAapLeuValThProAla 107
Db      588 AGAACAAATGATTGTGCGCATCAATCTTCAGCATATCAACAAAGACTGGTTACCCCGCT 529
Qy      108 A1Aa1Aa1ProG1nA1a1Val1ySerserSer1eAla1Aa1ProSerArProG1uPProAn 127
Db      528 GCGAGTCCCGCCCTCAGGAGGAGAGTCTCTATGCGAGCTCCCTGAGACCCGAGCCCAAT 469
Qy      128 ArgCyG1SerCyArG1yArG1yVal1G1yLeuThG1yPhe1ySArG1ySg1yAa1 147
Db      468 CGATGGGCAATCTTGACGAAAGCCTGTGATTCACAGATTTAAGTGTGCTGTGGCAAC 409
Qy      148 LeuTyTArG1A1eU1a1eU1a1rG1ySerArP1yH1eThrCyArThTyTArP1yT1yA1a 167
Db      408 CTCTACTGCGCTTATCATCGTACTCGACAACACACTTGACATATGACTACAAAGCC 349
Qy      168 A1aG1yG1nG1uA1a1e1a1yS1a1aAnProLeuVal1A1a1aG1u1ySVal1y1yS 187
Db      348 GGAAGGAGAAAGCGATGCGAAAGCTAACTCTTGCTGTGCGCGAGAAAGTTGTCAAG 289
Qy      188 Phe 188
Db      288 TTT 286

RESULT 3
LOCUS   BU592532/c      729 bp      mRNA      linear      EST 22-OCT-2003
DEFINITION   BU592532 normalized full length cDNA library, chloronemata, caulonemata and malformed buds Physcomitrella patens subsp. patens cDNA clone pphb6h03 3', mRNA sequence.
ACCESSION   BU592532
VERSION     BU592532.1 GI:37834520
KEYWORDS    EST.
SOURCE      Physcomitrella patens subsp. patens
ORGANISM    Physcomitrella patens subsp. patens
REFERENCE   1 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta; Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella.
AUTHORS     Nishiyama,T., Fujita,T., Shin-I,T., Seki,M., Nishide,H., Uchiyama,I., Kamiya,A., Carninci,P., Hayashizaki,Y., Shinozaki,K., Kohara,Y. and Haasbe,M.
TITLE       Comparative genomics of Physcomitrella patens gametophytic transcriptome and Arabidopsis thaliana: implication for land plant evolution
JOURNAL    Proc. Nect. Acad. Sci. U.S.A. 100 (13), 8007-8012 (2003)
PUBMED     12808149
COMMENT     Contact: Tadao Shii-1
              Center For Genetic Resource Information
              National Institute of Genetics
              111 Yata, Mishima, Shizuoka 411-8540, Japan
              Tel: 81-559-81-6856
              Fax: 81-559-81-6855
              Email: tshii@genes.nig.ac.jp
              A backbone of the vector is basically from pBluescript II (KS), that was in vivo excised from a 1-FLC phase vector (Carninci et al. 2001). 5' end of the cDNA that was digested with XhoI was ligated to SalI site of the vector and the 3' end including polyA tail was ligated to BamHI site of the vector(5'- ggaGAgAgAgATCCACCCCTGgAgAgATTTTTTTTTTTTAA-3' was used as a 1st 3' primer, and 5'-ggTCTCGATCATGCTGTTCAGACAGCATGATGATCGAGAACGNNNN-3' as 2nd 5'-hairpin primer, giving the following 5' boarder sequence, AGGCCAAATCGCGAGCTGATTCGCGAAGCCG). cDNA insert could be amplified with conventional T7 and T3 primers. This full-length cDNA library was generated according to the method described in Nishiyama et al. (2003).
              Protonemata were blended by the POLYTRON, and then cultivated on the BCDATG medium for 13-14 days under the continuous light. These clones are available from RIKEN Bio Resource Center (http://www.brc.riken.go.jp/lab/epd/Eng/index.html). The database

```

```

FEATURES
source
    of Physcomitrella EST clones is available at the PH
    (http://moses.nibb.ac.jp):
    Location/Qualifiers
        1..729
            /organism="Physcomitrella patens subsp. pa
            /mol_type="mRNA"
            /db_xref="taxon:145481"
            /clone="pphb6h03"
            /tissue_type="mixture of chloronemata, cau
            /clone.lib="normalized full length cDNA 1:
            chloronemata, caulonemata and malformed bu

ALIGNMENT Scores:
Pred. No.:      1,73e-70      Length:      729
Score:          834.00      Matches:      160
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:     84.76%      Indels:      0
DB:              Gaps:      0

US-10-716-089-20 (1-188) x BU592532 (1-729)

Qy      29 GlySerGlnAlaThmCtIleUcYsSerIySArG1yTArG1uThr 48
Db      728 GCGAGCCAAAGCTTACATGGGAGTTGTGCTCGAAGTCTACCGAGAGCAC 669
Qy      49 LySmetThrAlaLeuA1aG1uG1nA1aThrG1nA1a1Aa1aThr 68
Db      668 AAGATGACGGCTTTAGCTGACGAGACCACTCAGGCTGCTCAGAGCAAC 609
Qy      69 A1a1aVal1G1nProProAlaProVal1H1aG1uThr1ySLeuThrCySg 88
Db      608 GCTGCTGTTCAAGCCCCCGCTCCTGTGATACATGACCAAGCTCACAATGCC 549
Qy      89 ThrMetIleVal1ProH1gInSerSerSerTyTgInGlnAapLeuValT 108
Db      548 ACAATGATTGTGCCGATCAATCTTCAGCTATCAACAAGACTGTAA 489
Qy      109 A1a1aProG1nA1a1Val1ySerSer1eAla1Aa1Aa1ProSerArProG 128
Db      488 GCTGCCCTCAGGCGATGAAGTCTTATCGCAGCTCCCTCTAGACCCG 429
Qy      139 CySgIySerCyArG1yArG1yVal1G1yLeuThrG1yPhe1ySArG1 148
Db      428 TCGGATCTTGACAGAAAGGCTGTGATTCACAGATTTAAGTCCG 369
Qy      149 TyTArG1A1eU1a1eU1a1rG1ySerArP1yH1eThrCyArThTyTArP 168
Db      368 TACTCGCTTATCATCGTACTCGACAACACACTTGACATATGACT 309
Qy      169 G1yG1nG1uA1a1e1a1yS1a1aAnProLeuVal1A1a1aG1u1yS 188
Db      308 GGGCAGAAAGCATTCGAAAGCTAACTCTTGCTGTGCGCGAGAAAG 249

RESULT 4
LOCUS   BU592791/c      733 bp      mRNA      linear      EST 22-OCT-2003
DEFINITION   BU592791 normalized full length cDNA library, chloronemata and malformed buds Physcomitrella patens cDNA clone pphb6h21 3', mRNA sequence.
ACCESSION   BU592791
VERSION     BU592791.1 GI:37834779
KEYWORDS    EST.
SOURCE      Physcomitrella patens subsp. patens
ORGANISM    Physcomitrella patens subsp. patens
REFERENCE   1 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella.
AUTHORS     Nishiyama,T., Fujita,T., Shin-I,T., Seki,M., Nishide,H., Uchiyama,I., Kamiya,A., Carninci,P., Hayashizaki,Y.,

```

TITLE	Comparative genomics of <i>Physcomitrella patens</i> gametophytic transcriptome and <i>Arabidopsis thaliana</i> : implication for land plant evolution
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 100 (13), 8007-8012 (2003)
PUBMED	12808149
COMMENT	Contact: Tadao Shin-1

Location/Qualifiers

```

1. 733
/organism="Phycomitrella patens subsp. patens"
/mol_type="mRNA"
/sdb_species="patens"
/db_xref="taxon:145481"
/clone="ppb6h21"
/tissue_type="mixture of chloronemata, caulonemata and
malformed buds"
/clone_lib="normalized full length cDNA library,
chloronemata, caulonemata and malformed buds"

```

Alignment Scores:	
Pred. No.:	1,74e-70
Score:	834.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	84.76%
DB:	3
	Gaps: 0
	Matches: 733
	Mismatches: 160
	Conservative: 0
	Mismatches: 0
	Indels: 0
	Gaps: 0

US-10-716-089-20 (1-188) x BJ592791 (1-733)

QY	2	GlySerGlnAla1ThrMetGlyLeuGlySerLeuGlyCysTrpArg3LysIleValMetGlnAla	48
Db	731	GGCAGCGAAGCTACCAATGGGGTTGGCTCGAAGTCTCCAGAGACAGTATCAAGCG	6722
QY	49	LYSMetThrAlaLeuAlaGlnGlnAlaThrGlnAlaAlaGlnAlaThrSerAlaThrAla	68
Db	671	AAGATACCGCTTTAGCTGAGCAGACCACTCAGGCTGTCAGGGAGACATCTGCCACAGCT	6121
QY	69	AlaAlaValaGlnProProAlaProValHisGluThrLysLeuThrCysGluValaGlnArg	88
Db	611	GGCGCTGTTCAAGCCCCCGCTCTCTTAATGAGACCAAGCTCAATCGAGGTTGAGAGA	5522
QY	89	ThrMetIleValProHisGlnSerSerSerTrpGlnIleAspLeuValThrProAlaAla	108
Db	551	ACAATGATTGTGCCGATCAATCTTCACAGCTATCAACAAAGACCTGGTTATCCCCCGTCGCA	4922
QY	109	AlaAlaProGlnAlaValaLysSerSerIleAlaAlaProSerArgProGluProAsnArg	128

Db	491	GCTGCCCCCAGGACGATGAAGTCTCTATACCGACAGCTCCCTCTAGAACCCG
Qy	129	CysgIserCySarGLysarGLyValgIleuthnGlyPheLysCySarGc
Db	431	TGCGGATCTTCAGGAGAGCGGTGGATTGGACAGATTTTAAGTGTGCT
Qy	149	TyrCyAlaLeuHlaArgTyrSerAspLysHlaThrCysThrTyrAspT
Db	371	TACTGCGCTTAACTACGCGTACTCGGCAACAACACTGTGCACATATGACT
Qy	169	GLyngInGluAlaIleAlaLysAlaAsnProLeuValValAlaGluLysV
Db	311	GGGCGAGAGAGGATTGGCAAACTATCTCTTGTGTGTGGCCGAGAAAG

RESULT 5			
BJ584946/c			
LOCUS	BJ584946	747 bp	mRNA linear
DEFINITION	BJ584946 normalized full length cDNA library, chloroform ethanol-precipitated, ligated into pUC19, transformed into <i>Escherichia coli</i> DH5α, and sequenced. The sequence is identical to the sequence of the cDNA clone pphb26k21 3', mRNA sequence.		

ACCESSION	BJ584946
VERSION	
KEYWORDS	BJ584946.1 GI:37826934
SOURCE	EST.
ORGANISM	<i>Physcomitrella patens</i> subsp. <i>patens</i>
	<i>Physcomitrella patens</i> subsp. <i>patens</i>

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 100 (13), 8007-8012 (1993)
 PUBMED 12808149
 COMMENT Contact: Tadasu Shin-i

FEATURES	Location/Qualifiers
source	1. .747

```
1..74
/organism="Physcomitrella patens subsp. pa
/mol_type="mRNA"
/sub_species="patens"
/db_xref="taxon:145481"
/clone="hphb26K21"
/clone_type="mixture of chloronemata, cau
malformed bud"
/clone_11b="normalized full length cDNA 11
chloronemata, caulonemata and malformed
```

ORIGIN

Alignment Scores:

Pred. No.:	1.33e-69	Length:	747
Score:	825.00	Matches:	159
Percent Similarity:	98.76%	Conservative:	0
Best Local Similarity:	98.76%	Mismatches:	2
Query Match:	83.84%	Indels:	0
DB:	3	Gaps:	0

US-10-716-089-20 (1-188) x BU584946 (1-747)

```
OY 28 PheGlySerGlnAlaIleuValMetGlyLeuCySerIysCysTyrArgGluThrValMetGln 47
DB 747 TTCGGAGCCCAAGNTCCATGGGGTTGTGCTCGAAGTGTCTACCGAAGACAGTCATGCA 688
OY 48 AlaIysMetThrAlaIleuValaGluGlnAlaThrGlnAlaIleuValaIleuValaThr 67
DB 687 GCGAAGATGACGGCTTTAGCTGAGCAAGCGACTGAGCTGCTCAGCGACATCTGCCACA 628
OY 68 AlaAlaIleuValaGlnProProAlaProValaIleuValaIleuValaIleuValaIleu 87
DB 627 GCTGCTGCTGTTCAAGCCCCCGCTCTCTGTAACATGAGACCAAGCTCATGCGAGGTTGAG 568
OY 88 ArgThrMetIleValaProIleGlnSerSerSerTyrGlnGlnAspLeuValaThrProAla 107
DB 567 AACAACATGATTTGTGCGCATCATCTTCAGCTATCAACAGACCTGGTTACCCCCGCT 508
OY 108 AlaAlaIleuValaIleuValaIleuValaIleuValaIleuValaIleuValaIleuVala 127
DB 507 GGAAGTGGCCCTCAAGCACTGAAGTCTCTATCGAGCTCCCTTGAACCCGACCAT 448
OY 128 ArgCySerGlySerCyArgGlyArgValaGlyLeuThrGlyPheIysCysArgCysGln 147
DB 447 CGATCGCGATCTTGGCAGAAAGCGCTGGATTGACAGATTTAAGTGTGCTGGCAAC 388
OY 148 LeuTyrCysAlaIleuValaIleuValaIleuValaIleuValaIleuValaIleuVala 167
DB 387 CTCCTACTGGCTTTTACATGCTGCTGCAACAACACTTACACATATACCAAGCC 328
OY 168 AlaGlyGlnGlnAlaIleuValaIleuValaIleuValaIleuValaIleuValaIleuVala 187
DB 327 GCAAGGCGAAGAACCATTCGAAAGCTATCTCTTGTGTGCGCAGAGAGTTGTGCAAG 268
OY 188 Phe 188
DB 267 TTT 265
```

RESULT 6
BU585055/c 727 bp mRNA linear EST 22-OCT-2003
LOCUS BU585055 normalized full length cDNA library, chloronemata, 727 bp
DEFINITION BU585055 normalized full length cDNA library, chloronemata, 727 bp
cautionemata and malformed buds Physcomitrella patens subsp. patens
cDNA clone pphb26n05 3', mRNA sequence.

ACCESSION BU585055
VERSION BU585055.1 GI:37827043
KEYWORDS EST.
SOURCE Physcomitrella patens subsp. patens
ORGANISM Physcomitrella patens subsp. patens
REFERENCE 1 (bases 1 to 727)
AUTHORS Nishiyama, T., Fujita, T., Shin-I, T., Seki, M., Nishide, H.,
Uchiyama, I., Kamiya, A., Carninci, P., Hayashizaki, Y., Shinozaki, K.,
Kohara, Y. and Hasebe, M.
TITLE Comparative genomics of Physcomitrella patens gametophytic
transcriptome and Arabidopsis thaliana: implication for land plant
evolution
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 100 (13), 8007-8012 (2003)
PUBMED 12808149
COMMENT Contact: Tadao Shin-I
Center For Genetic Resource Information
National Institute of Genetics

FEATURES

source

1111 Yata, Miehima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp
A backbone of the vector is basically from pBluesc
that was in vivo excised from a 1-F1C phage vector
2001). 5' end of the cDNA that was digested with X
to SalI site of the vector and the 3' end including
ligated to BamHI site of the
vector(5'- gAAGAGAGAGAGATCCACCCCTGAGAGATTGTTTTT
used as a 1st 3' primer, and
5'-ggTCTCGATGATCGCTGTTCACAGACGATGATCTGACACCGN
5'-hairpin primer, giving the following 5' boarder
AGGCCAATGCGCCGACGCTGATTCCTCGAAGACCG). cDNA inate
amplified with conventional T7 and T3 primers. This
cDNA library was generated according to the method
Nishiyama et al. (2003).
Protonemata were blended by the POLYTRON, and then
The BODATG medium for 13-14 days under the contin
These clones are available from RIKEN Bio Resource
(http://www.brc.riken.go.jp/lab/epd/bmg/index.html)
of Physcomitrella EST clones is available at the Ph
(http://mos.nibb.ac.jp).

location/Qualifiers

1.727
/organism="Physcomitrella patens subsp. pa
/mol_type="mRNA"
/sub_species="patens"
/db_xref="taxon:145461"
/clone="pphb26n05"
/rissue_type="mixture of chloronemata, cau
/clone_lib="normalized full length cDNA l
chloronemata, cautionemata and malformed bu

ORIGIN

Alignment Scores:

Pred. No.:	1.6e-69	Length:	727
Score:	824.00	Matches:	159
Percent Similarity:	99.38%	Conservative:	0
Best Local Similarity:	99.38%	Mismatches:	1
Query Match:	83.74%	Indels:	0
DB:	3	Gaps:	0

US-10-716-089-20 (1-188) x BU585055 (1-727)

```
OY 29 GlySerGlnAlaIleuValMetGlyLeuCySerIysCysTyrArgGluThrVal 48
DB 726 GCGAGCCCAAGCTACCAATGGGGTTGTGCTCGAAGTGTCTACCGAAGACAG 667
OY 49 LysMetThrAlaIleuValaGluGlnAlaThrGlnAlaIleuValaIleuValaIleuVala 68
DB 666 AAGATGACGGCTTTAGCTGAGCAACCACTCAAGCTGCTCAGGCGCACT 607
OY 69 AlaAlaIleuValaIleuValaIleuValaIleuValaIleuValaIleuValaIleuVala 88
DB 606 GCTGCTGTTACGCCCCCGCTCTCTTACATGAGCAAGCTCACAATGCG 547
OY 89 ThrMetIleuValaProIleGlnSerSerSerTyrGlnGlnAspLeuValaThr 108
DB 546 ACAAATGATTTGCGCCATCAATCTTCAGCTATCAACAAGACCTGGTTA 487
OY 109 AlaAlaProGlnAlaIleuValaIleuValaIleuValaIleuValaIleuValaIleuVala 128
DB 486 GCTGCCCTCAGGACGAGTAAAGCTCTATCGAGCTCCTCTAGACCGG 427
OY 129 CysGlySerCyArgGlyArgValaGlyLeuThrGlyPheIysCysArgCys 148
DB 426 TCGCGATCTTCGAGGAAAGCGTTGGATTGACAGGATTTAAGTGTGCGCA 367
OY 149 TyrCysAlaIleuValaIleuValaIleuValaIleuValaIleuValaIleuValaIleuVala 168
DB 366 TACTGCGCTTACATCGGTACTCGGACAAACACACTTGCACATATGACT 307
```

```

y 169 GYGVINGTUAIAIEALYALYALAspProleuVA1A1AAGUlyeVal1ygshe 188
Db 306 GGGCGAGGAGCGATTGCGAAGCTATCTCTTGTGTGGCCGAGAGGTTGTCAAGTTT 247

RESULT 7
BU585014/c
LOCUS
DEFINITION
BU585014 721 bp mRNA linear EST 22-OCT-2002
BU585014 normalized full length cDNA library, chloronemata,
cDNA clone pphb26904 3', mRNA sequence.
ACCESSION
BU585014
VERSION
BU585014.1 GI:37827002
SOURCE
EST.
ORGANISM
Physcomitrella patens subsp. patens
Physcomitrella patens subsp. patens
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella.
AUTHORS
Nishiyama,T., Fujita,T., Shin-i,T., Seki,M., Nishide,H.,
Uchiyama,T., Kamaya,A., Carninci,P., Hayashizaki,Y., Shinozaki,K.,
Kohara,Y. and Haegebe,M.
Comparative genomics of Physcomitrella patens gametophytic
transcriptome and Arabidopsis thaliana: implication for land plant
evolution
JOURNAL
Proc. Natl. Acad. Sci. U.S.A. 100 (13), 8007-8012 (2003)
PUBMED
12808149
COMMENT
Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tehin@genes.nig.ac.jp
A backbone of the vector is basically from pBluescript II (KS),
that was in vivo excised from a 1-flCt phage vector (Carninci et al
2001). 5' end of the cDNA that was fused with XhoI was ligated
to SalI site of the vector and the 3' end including polyA tail was
ligated to BamHI site of the
vector(5'- gGAGAGAGAGGATCCAAACCCGgAgATTGTTTTTTTTTTTAA-3' was
used as a 1st 3' primer, and
5'-gGTTCTGATCATCGTGTCTGTCGAGCGATCGAGACTCGAGAACGNNNN-3' as 2nd
5'-hairpin primer, giving the following 5' boarder sequence,
AGGCCAATATCGCGAGCTCGAATTCGTCAAGACG). cDNA insert could be
amplified with conventional T7 and T3 primers. This full-length
cDNA library was generated according to the method described in
Nishiyama et al. (2003).
Protonemata were blended by the POLYTRON, and then cultivated on
the PCDATG medium for 13- 14 days under the continuous light.
These clones are available for RIKEN Bio Resource Center
(http://www.brc.riken.go.jp/labb/epd/Bng/index.html). The database
of Physcomitrella EST clones is available at the PHYSCOBASE
(http://mos.nibb.ac.jp).
Location/Qualifiers
1. 721
/organism="Physcomitrella patens subsp. patens"
/mol_type="mRNA"
/sub_species="patens"
/db_xref="taxon:145481"
/clone="pphb26904"
/tissue_type="mixture of chloronemata, caulonemata and
malformed buds"
/clone_lib="normalized full length cDNA library,
chloronemata, caulonemata and malformed buds"

ORIGIN
Alignment Scores:
Pred. No.: 1,386-67 Length: 721
Score: 804.00 Matches: 155
Percent Similarity: 98.73% Conservative: 0
Best Local Similarity: 98.73% Mismatches: 2
Query Match: 81.71% Indels: 0
Gaps: 3

```

US-10-716-089-20 (1-188) x BU585014 (1-721)
Oy 32 A1athmctGlyLeuCysserlyScyTYrAagluThrValmctGln
Db 720 GCTACCAAGGGGGTGTGTCTGTAAGAGNTACCGAGAACACTCATGCAAG
Oy 52 A1aleuA1agluGlnA1athrGlnA1a1aglnA1athrSerA1aThrA
Db 660 GCTTTAGCTGAGCAAGCACCCTAGGCTGCTCAAGCGACACTTGTCCACAG
Oy 72 GlnProProlaProValHisgluThrlyLeuThrCysegluValGlnA
Db 600 CAGCCCCCGCTCTGTACATGAGACCAAGCTCATGCGAGGTGAGA
Oy 92 ValProHisGlnSerSerSerTyrglnGlnAspleuValThrProAla
Db 540 GTGCCGACATCATCTTCCAGCTATCAACAAAGACTGGTTACCCCGCTG
Oy 112 GlnA1aVallySerSerSerleuA1a1aProSerArgProgluProAsnA
Db 480 CAGGCAAGTGAAGTCTCTTACGAGCTCCCTCTAGACCCGAGCCCAATC
Oy 132 CySarGlySarGValGlyLeuThrglyPheLyCyargCysegluAsnL
Db 420 TGCAGGAAGCGGTGGATTGATGACAGGATTTAAGTGCCTGTGGCAAC
Oy 152 LeuHisAArgTySerAspLyshisThrCyserThrTyAspTytyrlyAlaA
Db 360 TTACTATCGGTACTCGACAAACACTTGCACATTAAGCTACCAAGCGC
Oy 172 A1a1leuA1a1aAsnProleuValA1a1agluValVallyLysP
Db 300 CGGATTCGGAAGCTATCTCTTGTCTGTGCGCGAGAAAGTTGTCAAGT
RESULT 8
LOCUS BU606134/c
DEFINITION BU606134 normalized full length cDNA library, chlor
LOCUS BU606134
VERSION BU606134.1 GI:37848126
KEYWORDS EST.
SOURCE Physcomitrella patens subsp. patens
ORGANISM Physcomitrella patens subsp. patens
REFERENCE Bryopsida; Vitidiplantae; Streptophyta; Embryophyta
AUTHORS Bryopsida; Funariidae; Funariales; Funariaceae; Phy-
1 (baes 1 to 702)
Ushiyama,T., Fujita,T., Shin-I,T., Seki,M., Nishid
Uchiyama,T., Kamiya,A., Carninci,P., Hayashizaki,Y.
Kohara,Y. and Hasebe,M.
Comparative genomics of Physcomitrella patens gemnetic
transcriptome and Arabidopsis thaliana: Implication
evolution
Proc. Natl. Acad. Sci. U.S.A. 100 (13), 8007-8012 (
12808149
CONTACT: Tadao Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel.: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshin@ggenes.nig.ac.jp
A backbone of the vector is baically from pBluescr
that was in vivo excised from a 1-PLC phage vector
2001). 5' end of the cDNA that was digested with Xh
to SalI site of the vector and the 3' end including
ligated to BamHI site of the
vector(5'- gaagagagagatccacacctgagagtttttttttttt
used as a 1st 3' primer, and
5'-ggtttcgagatcgttcgttcgagacgcatgactgagAACgann
5'-hairpin primer, giving the following 5' boarder

51	
661	
71	
601	
91	
541	
111	
481	
131	
421	
151	
361	
171	
301	
2003	
al.	
ed	
was	
is	
id	

AGCCCAATCCGCCAGCTCGAATTCGTGAGAACCG). cDNA insert could be amplified with conventional T7 and T3 primers. This full-length cDNA library was generated according to the method described in Nishiyama et al. (2003). Protonemata were blended by the POLYTRON, and then cultivated on the BCDATG medium for 13-14 days under the continuous light. These clones are available from RIKEN Bio Resource Center (<http://www.brc.riken.go.jp/lab/epd/Eng/index.html>). The database of Physcomitrella EST clones is available at the PHSCODbase (<http://mos.nibb.ac.jp>).

FEATURES

source

1..702
Location/Qualifiers
/organism="Physcomitrella patens subsp. patens"
/mol_type="mRNA"
/sub_species="patens"
/db_xref="taxon:145481"
/clone="pPhn35f21"
/tissue_type="mixture of chloronemata, caulonemata and rhizoid-like protonemata"
/clone_1lb="normalized full length cDNA library, chloronemata, caulonemata and rhizoid-like protonemata"

ORIGIN

Alignment Scores:

Pred. No.: 7,396-66 Length: 702
Score: 786.00 Matches: 151
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 79.88% Indels: 0
DB: 3 Gaps: 0

US-10-716-089-20 (1-188) x BU606134 (1-702)

OY 38 SerLySVYrYrAGlUThrVAlMeCglnAlaLyMeCThrAlaLeuAlaGlUglnAla 57
DB 702 TCGAAGTGTACCGAGAGACAGTCAATGACGAGAGATGACGCTTTAGCTGAGCAAGCC 643
OY 58 ThrGlnAlaAlaGlAlaThrSerAlaThrAlaAlaAlaAlaGlnProProAlaProVal 77
DB 642 ACTCAGGCTGCTCAGGAGCATCTCCACAGCTGCTGCTTCCAGCCCTCTCTGTA 583
OY 78 HlsglUThrLyLeuThrCySglUValGluArGThrMetIleValProHlsglInsSer 97
DB 582 CATGAGACCAAGCTCAATGCAAGTTGAGAGAACATGATTTGCTCCGATCAATCTCC 523
OY 98 SerTyrgInglInAspLeuValThrProAlaAlaAlaAlaProGlnAlaValLySerSer 117
DB 522 AGCTATCAACAAGACTGTGTTACCCCGCTGCACTGCCCTCAGGCAGTGAAGCTCT 463
OY 118 ILeAlaAlaProSerArGProGluProAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 137
DB 462 ATCGAGCTCCCTCTAGACCCGAGCCCAATCGATCTTTCAGGAGAGCTTTGCA 403
OY 138 LeuThrGlyPheLyCySArGCGlUThrValLeuValLeuAlaAlaAlaAlaAlaAla 157
DB 402 TTGACAGATTTAGTGTGCTGTGCAACCTCTTCTGCTTTCATGCGGACTCGAAC 343
OY 158 LySAlaThrCySArGThrTyAspTyTrLyAlaAlaAlaGlynglnAlaAlaAlaAlaAla 177
DB 342 AAACACACTTCACATATGACTACAAAGCCGAGGAGGAGAAAGCATTTGCGAAAGCTAAT 283
OY 178 ProLeuValValAlaAlaGluLyValValLySArPhe 188
DB 282 CCTCTGTGTGCGCGAGAGAGTGTCAAGTTT 250

RESULT 9
BU591761/c 704 bp mRNA linear EST 22-OCT-2003
LOCUS BU591761 normalized full length cDNA library, chloronemata,
DEFINITION caulonemata and malformed buds Physcomitrella patens subsp. patens
cDNA clone pphb44121 3', mRNA sequence.
ACCESSION BU591761
VERSION BU591761.1 GI:37833749

KEYWORDS

SOURCE

EST.
Physcomitrella patens subsp. patens
Physcomitrella patens subsp. patens

ORGANISM

Bryophyta; Viridiplantae; Streptophyta; Embryophyta
Bryopsida; Funariidae; Funariales; Funariaceae; Phy

REFERENCE

1 (bases 1 to 704)
Nishiyama, T., Fujita, T., Shin-I, T., Seki, M., Nishio,
Uchiyama, I., Kamiya, A., Carninci, P., Hayashizaki, Y.,
Kohara, Y. and Hasebe, M.
Comparative genomics of Physcomitrella patens gametophyte
transcriptome and Arabidopsis thaliana: implications for
evolution

TITLE

Proc. Natl. Acad. Sci. U.S.A. 100 (13), 8007-8012

JOURNAL

Proc. Natl. Acad. Sci. U.S.A. 100 (13), 8007-8012

COMMENT

Contact: Tadao Shin-I
Center For Genetic Resource Information
National Institute of Genetics
111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp

A backbone of the vector is basically from pBluescript
that was in vivo excised from a 1-FLC phage vector
2001). 5' end of the cDNA that was digested with XbaI
to SalI site of the vector and the 3' end including
ligated to BamHI site of the
vector(5'- gGAGAGAGAGATCCACCTGgAGATTTTCTTTTCTT
used as a 1st 3' primer, and
5'-gGTTCTGAGTCAATCGCTGTCAGACGAGTATCTGAGAACCGN
5'-hairpin primer, giving the following 5' boarder
AGCCCAATCCGCCAGCTCGAATTCGTGAGAACCG). cDNA insert
amplified with conventional T7 and T3 primers. This
cDNA library was generated according to the method
Nishiyama et al. (2003).
Protonemata were blended by the POLYTRON, and then
These clones are available from RIKEN Bio Resource
(<http://www.brc.riken.go.jp/lab/epd/Eng/index.html>)
of Physcomitrella EST clones is available at the PH
(<http://mos.nibb.ac.jp>).

FEATURES

source

1..704
Location/Qualifiers
/organism="Physcomitrella patens subsp. pa
/mol_type="mRNA"
/sub_species="patens"
/db_xref="taxon:145481"
/clone="pPhb44121"
/tissue_type="mixture of chloronemata, cau
malformed buds"
/clone_1lb="normalized full length cDNA l
chloronemata, caulonemata and malformed bu

ORIGIN

Alignment Scores:

Pred. No.: 5,536-65 Length: 704
Score: 777.00 Matches: 149
Percent Similarity: 99.33% Conservative: 0
Best Local Similarity: 99.33% Mismatches: 1
Query Match: 78.96% Indels: 0
DB: 3 Gaps: 0

US-10-716-089-20 (1-188) x BU591761 (1-704)

OY 39 LySCySArYrAGlUThrVAlMeCglnAlaLyMeCThrAlaLeuAlaAla 58
DB 702 AAGTCTACCGAGAGACAGTCAATGCAAGAGATGAGCGCTTTAGCT 643
OY 59 GlnAlaAlaGlnAlaThrSerAlaThrAlaAlaAlaAlaGlnProPro 78
DB 642 CAGCTGCTCAGGAGACATCTGCAACAGCTCTGCTGTTACGCCCCC 583
OY 79 GluThrLySLeuThrCySglUValGluArGThrMetIleValProHlsg 98

Db 582 GAGACCAAGCTCACATCGAGGTGAGAGAACATGATGTCGCCATCAATCTTCCAGC 523

Qy 99 TGTGNGINAPLEUVALThrProAlaAlaAlaProGlnAlaValIysSerSerIle 118

Db 522 TATCAACAAGACTGTGTAACCCCGCTGAGCTGCCCTCAGACGATGAAGCTCTATC 463

Qy 119 AlaAlaProSerArgProGluProAlaAlaArgCysGlySerCysArgIysArgValGlyLeu 138

Db 462 GCAGTCCCTCTAGACCCGAGCCCAATCATGATCGGATCTTGACGAGAACGCTTGATG 403

Qy 139 ThrGlyPheIysCysArgCysGlyValenLeuTyrcysAlaLeuHisArgTySerAspIys 158

Db 402 ACAGATTATTAAGTGTGCTGTGGCAACCTCTACTCGCTTACATCGGTACTCGACAAA 343

Qy 159 HisThrCysThrTyAspTyrcysAlaIaIaGlyGlnGlnAlaIaAlaIysAlaAsnPro 178

Db 342 CACACTTCNCATATGACTACAAAGCCGAGGGCAGAAAGCATTCGAAAGCTATCT 283

Qy 179 LeuValValAlaGluIysValValIysPhe 188

Db 282 CTGTGCGTGGCCGAGAGGTTCTCAAGTTT 253

RESULT 10

LOCUS BU941820 719 bp mRNA linear EST 13-JUN-2005

DEFINITION BU941820 phf1 full-length cDNA library Physcomitrella patens subsp.

ACCESSION BU941820 GI:67568996

VERSION BU941820.1 GI:67568996

KEYWORDS EST.

SOURCE Physcomitrella patens subsp. patens

ORGANISM Physcomitrella patens subsp. patens

REFERENCE 1 (bases 1 to 719)

AUTHORS Fujita,T., Nishiyama,T., Shin-I,T., Kohara,Y. and Hasebe,M.

TITLE Physcomitrella patens EST at a stage of the first asymmetric cell division of protoplasts

JOURNAL Unpublished (2005)

COMMENT Contact: Tadasu Shin-I
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshin@genes.nig.ac.jp

Protonemata were inoculated on BCDATG medium for every ca. 5 days. Protonemata were isolated from the protonemata, further incubated at 25C under continuous light for 2-3 days. The regenerated cells, which were rich in cells at a stage during the first asymmetric cell division, were collected. Total RNA was extracted for constructing a full-length cDNA library. The database of the EST clones is available at the PHYSCODbase (http://moss.nibb.ac.jp).

FEATURES

source

1..719

location/Qualifiers

1..719

organism="Physcomitrella patens subsp. patens"

mol_type="mRNA"

sub_species="patens"

db_xref="Caxon:145481"

clone="phf16106"

issue_type="regenerated protoplasts (chloronemata)"

dev_stage="at the first asymmetric cell division of protoplasts"

clone_lib="phf1 full-length cDNA library"

note="Protonemata were inoculated on BCDATG medium for every ca. 5 days. Protoplasts were isolated from the protonemata, further incubated at 25C under continuous light for 2-3 days. The regenerated cells, which were rich in cells at a stage during the first asymmetric cell division, were collected. Total RNA was extracted for constructing a full-length cDNA library."

Alignment Scores:

Pred. No.: 2,78e-58 Length: 719

Score: 708.00 Matches: 132

Percent Similarity: 83.15% Conservative: 16

Best Local Similarity: 74.16% Mismatches: 30

Query Match: 71.95% Indels: 0

DB: 3 Gaps: 0

US-10-716-089-20 (1-188) x BU941820 (1-719)

Qy 11 ThSerGlnAlaProGluGlyProValMetCysIysAsnLeuCysGlyP 30

Db 3 ACCGCCAGACCCCGAGGGGCTGTGATGTGAAGACGCTGGGTTT 62

Qy 31 GlnAlaThrMetGlyLeuCysSerIysCysArgTyrcysGluThrValMetG 50

Db 63 CAGGCACTATGGTCTCTGCTCGAAGCTTACAGGAGACCGTCATTC 122

Qy 51 ThrAlaLeuAlaGluGlnAlaThrGlnAlaAlaGlnAlaThrSerAlaT 70

Db 123 ACCGCGTGGCCGAGCAGGCGCACTCAAGCCGCTCAGGTGCGATCCG 182

Qy 71 ValGlnProProAlaProValHisGluThrIysLeuThrCysGlyValG 90

Db 183 GCTCAGCCTCCAGTTCCTTATGGAAGAGCACAAGAGCTCATTTAGGCCG 242

Qy 91 IleValProHisGlnSerSerSerTyrcGlnGlnAspLeuValThrProA 110

Db 243 ATCAACCCGCTCAATCTTCCAGCCACCATCCAGTTGAGGTTGCTCCGG 302

Qy 111 ProGlnAlaValIysSerSerIleAlaAlaProSerArgProGluProA 130

Db 303 CCCAGGTTGATGTTAGTCCCGCGTTGCCACACCTCCACACCGGACCAA 362

Qy 131 SerCysArgIysArgValGlyLeuThrGlyPheIysCysArgCysGlyA 150

Db 363 TCTTCAGAGAAAGCGGTGGATTGACAGATTTCAAGTTCCTGTGGGCG 422

Qy 151 AlaLeuHisArgTyrcysAspIysHisThrCysThrTyrcysAspTyrcysA 170

Db 423 GCTTACATCGGTATTCGACAAACATATGTTGCACTTATGATTAAGG 482

Qy 171 GluAlaIleAlaIysAlaAsnProLeuValAlaAlaGluIysValValI 1411

Db 483 GAGGCTATTGCAAAGGCTAATCTTTAGTTGTGGCCGAGAAAGTTGTTA 1411

RESULT 11

LOCUS BU165418 664 bp mRNA linear 2003

DEFINITION BU165418 full length cDNA library, chloronemata and gametophores Physcomitrella patens subsp. patens cd 1411

ACCESSION BU165418

VERSION BU165418.1 GI:18333403

KEYWORDS EST.

SOURCE Physcomitrella patens subsp. patens

ORGANISM Physcomitrella patens subsp. patens

REFERENCE 1 (bases 1 to 664)

AUTHORS Nishiyama,T., Fujita,T., Shin-I,T., Seki,M., Nishida, Uchiyama,I., Kamiya,A., Carninci,P., Hayashizaki,Y., Kohara,Y. and Hasebe,M.

TITLE Comparative genomics of Physcomitrella patens gametophores and Arabidopsis thaliana: implication for evolution

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 100 (13), 8007-8012 (2003)

PUBMED 12808149

COMMENT Contact: Tadasu Shin-I
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856

Fax: 81-559-81-6855

Email: tshnigenes.nig.ac.jp

A backbone of the vector is basically from pBluescript II (KS), that was in vivo excised from a 1-FLC phage vector (Carninci et al. 2001). 5' end of the cDNA that was digested with XhoI was ligated to SalI site of the vector and the 3' end including polyA tail was ligated to BamHI site of the vector (5' - gAAGAGAGAGATCCACCTGGAGAGTTTCTTTTCTTTTNN-3' was used as a 1st 3' primer, and 5'-ggTTCGAGTCATCGCTTCCTCCAGACGCGATGACTGAGAACGNNNN-3' as 2nd 5'-hairpin primer, giving the following 5' boarder sequence, AGGCCAATCGCCGACGATCGAATTCGAGAACCG). cDNA insert could be amplified with conventional T7 and T3 primers. This full-length cDNA library was generated according to the method described in Nishiyama et al. (2003). Protonemata were blended by the POLYTRON, and then cultivated on the BCDATG medium for 13-14 days under the continuous light. These clones are available from RIKEN Bio Resource Center (<http://www.brc.riken.go.jp/lab/epd/Eng/index.html>). The database of Physcomitrella EST clones is available at the PHYSCObase (<http://mos.nibb.ac.jp>)

FEATURES

SOURCE

1..664
Location/Qualifiers
/organism="Physcomitrella patens subsp. patens"
/mol_type="mRNA"
/sub_species="patens"
/db_xref="taxon:145481"
/clone="pPhac11"
/tissue_type="mixture of chloronemata and young gametophores with 2 to 5 leaves"
/clone_lib="full length cDNA library, chloronemata and young gametophores"

ORIGIN

Alignment Scores:

Pred. No.: 2,72e-56 Length: 664
Score: 687.00 Matches: 131
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 69.82% Indels: 0
DB: 3 Gaps: 0

US-10-716-089-20 (1-188) x BU165418 (1-664)

58 ThrGlnAlaAlaGlnAlaThrsGlnAlaAlaAlaValAlaGlnProAlaProVal 77
663 ACTCAGGCGCTCCAGCGCGACATCTGCGACAGCTGCTGCTCAGCCCCCTCCTGTA 604
78 HtAGTlThrLysLeuThrCyAGlUValGluArgThrMetIleValProHtAGlInsSer 97
603 CATGAGACCAACCTCATGATGAGGTTGAGAAACATGATGTCGCGATCATCTTCC 544
98 SerTyrGlnGlnAspLeuValThrProAlaAlaAlaAlaProGlnAlaValIysSerSer 117
543 AGCTATCAACAAAGACTGGTTACCCCGCTGACGCGCCCTCAGGCACTGAAGTCTCT 484
118 IleAlaAlaPheSerThrArgProGluProAlaAlaGlySerCyAGlUValGlyValGly 137
483 ATCGGAGCTCCCTCTAGAACCCGAGCCCAATCGCTGCGATCTTGGCAGGAACGCTTGG 424
138 LeuThrGlyPheLysCyAGlUValGlyAsnLeuValCysAlaLeuHtAspThrSerAsp 157
423 TTGACAGGATTTAAAGTGTGCTGTGGCAACCTCTACTGCGCTTTCATGCGGTACTCGAGC 364
158 LysHtAspThrCyAGlUValGlyAsnLeuValGlyValGlyValGlyValGlyValGly 177
363 AAACACACTTGCACATATGATCAAAAGCGCAGCGGCAAGGCAATTCGGAAGCTAAT 304
178 ProLeuValValAlaGlyValGlyValGlyValGlyValGlyValGlyValGlyValGly 188
303 CCTTTGTCTGTGCGCGAGAAAGGTGTCAAGTTT 271

RESULT 12

BU605573/c

LOCUS

BU605573 738 bp mRNA linear

DEFINITION

BU605573 normalized full length cDNA library, chloronemata and rhizoid-like protonemata Physcomitrella patens subsp. patens cDNA clone phn34m03 3', mRNA sequen-

ACCESSION

BU605573

VERSION

BU605573.1 GI:37847565

KEYWORDS

EST

SOURCE

Physcomitrella patens subsp. patens

ORGANISM

Physcomitrella patens subsp. patens

REFERENCE

Uchiyama, T., Fujita, T., Shin-I, T., Seki, M., Nishiyama, T., Kamiya, A., Carninci, P., Hayashizaki, Y., Kohara, Y. and Hasebe, M.

AUTHORS

Comparative genomics of Physcomitrella patens gametophore and Arabidopsis thaliana: implication in evolution

JOURNAL

Proc. Natl. Acad. Sci. U.S.A. 100 (13), 8007-8012 (2003)

PUBMED

12808149

COMMENT

CONTACT

Contact: Tadao Shin-I

Center For Genetic Resource Information

National Institute of Genetics

111 Yata, Mishima, Shizuoka 411-8540, Japan

Tel: 81-559-81-6856

Fax: 81-559-81-6855

Email: tshnigenes.nig.ac.jp

A backbone of the vector is basically from pBluescript II (KS), that was in vivo excised from a 1-FLC phage vector (Carninci et al. 2001). 5' end of the cDNA that was digested with XhoI was ligated to SalI site of the vector and the 3' end including polyA tail was ligated to BamHI site of the vector (5' - gAAGAGAGAGATCCACCTGGAGAGTTTCTTTTCTTTTNN-3' was used as a 1st 3' primer, and 5'-ggTTCGAGTCATCGCTTCCTCCAGACGCGATGACTGAGAACGNNNN-3' as 2nd 5'-hairpin primer, giving the following 5' boarder sequence, AGGCCAATCGCCGACGATCGAATTCGAGAACCG). cDNA insert could be amplified with conventional T7 and T3 primers. This cDNA library was generated according to the method described in Nishiyama et al. (2003). Protonemata were blended by the POLYTRON, and then cultivated on the BCDATG medium for 13-14 days under the continuous light. These clones are available from RIKEN Bio Resource Center (<http://www.brc.riken.go.jp/lab/epd/Eng/index.html>). The database of Physcomitrella EST clones is available at the PHYSCObase (<http://mos.nibb.ac.jp>)

FEATURES

SOURCE

1..738
Location/Qualifiers
/organism="Physcomitrella patens subsp. patens"
/mol_type="mRNA"
/sub_species="patens"
/db_xref="taxon:145481"
/clone="pPhn34m03"
/tissue_type="mixture of chloronemata and rhizoid-like protonemata"
/clone_lib="normalized full length cDNA library, chloronemata and rhizoid-like protonemata, caulonemata and rhizoid-like"

ORIGIN

Alignment Scores:
Pred. No.: 9,55e-56 Length: 738
Score: 682.00 Matches: 127
Percent Similarity: 83.14% Conservative: 16
Best Local Similarity: 73.84% Mismatches: 29
Query Match: 69.31% Indels: 0
DB: 3 Gaps: 0

US-10-716-089-20 (1-188) x BU605573 (1-738)

17 GlyProValMetCysLysAsnLeuCyAGlUValGlyPhePheGlySerGlnAlaAla 36
730 GGCCTGTGATGTGTGAAGACGTCGTGCGGTTCTTTGGCAGCGCAAGCCCA 671

QY 37 CysSerLysCySyrArGluThrValMetGlnAlaLysMetThrAlaLeuAlaGln 56
 Db 670 TGCTGAAGTGTATACAGGAGACCGCATGACAGGAAAGATATACCGCGTGCAGACG 611
 QY 57 AATrGlnAlaAlaGlnAlaThrSerAlaThrAlaAlaValAlaGlnProProAlaPro 76
 Db 610 GCCACTCAAGCGCGTCAAGGTGTGCATCCGCTGCATCCAGTGCATCCAGCTCACTTCT 551
 QY 77 ValHisGluThrLysLeuThrCysGluValAlaGluThrMetIleValProHisGlnSer 96
 Db 550 ATGGAAGAGGACAAAGAGCTATTGAGCGCCGACACATGTATACAAACCGCTCAACT 491
 QY 97 SerSerTyrrGlnGlnAspLeuValThrProAlaAlaAlaProGlnAlaValLysSer 116
 Db 490 TCCACCCACCATCCAGTTCAGGTCTCCGTTACAGTACGAGCCCGAGGTGTATGCTGCC 431
 QY 117 SerIleAlaAlaProSerArGProGluProAlaAlaGlySerCysAlaGlyVal 136
 Db 430 CCCGTGCGACACCTCCAGACCGGACCAACCGGTGGATCTTGACGAGAGCGCGT 371
 QY 137 GlyLeuThrGlyPheLysCysArGlyAsnLeuTyrrCysAlaLeuHisArGlySer 156
 Db 370 GGATTGACAGGATTCAGTGTCTGTGGGCACTTTTGTGCTTTCATCGGTATCT 311
 QY 157 AspLysHisIleThrCysThrTyrrAspTyrrLysAlaAlaGlnGlnAlaIleAlaLysAla 176
 Db 310 GACAAACATGATGACATTATGATTAACAAGCTGTGGACGAGAGCTATTGCACAAAGCT 251
 QY 177 AsnProLeuValAlaAlaGluLysValAlaLysPhe 188
 Db 250 AATCTTATAGTTGTGTGCCGAGAAAGTTGTTAAGTTC 215
 RESULT 13
 BJS98461/c
 LOCUS BJS98461 723 bp mRNA linear EST 22-OCT-2003
 DEFINITION BJS98461 normalized full length cDNA library, chloronemata, caulonemata and rhizoid-like protonemata Physcomitrella patens subsp. patens cDNA clone pphn23j11 3', mRNA sequence.
 ACCESSION BJS98461 GI:37840453
 VERSION BJS98461.1
 KEYWORDS EST.
 SOURCE Physcomitrella patens subsp. patens
 ORGANISM Physcomitrella patens subsp. patens
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta; Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella.
 1 (bases 1 to 723)
 Nishiyama,T., Fujita,T., Shin-I,T., Seki,M., Nishide,H., Uchiyama,T., Kamiya,A., Carninci,P., Hayashizaki,Y., Shinozaki,K., Kohara,Y. and Hasebe,M.
 Comparative genomes of Physcomitrella patens gametophytic transcriptome and Arabidopsis thaliana: Implication for land plant evolution
 Proc. Natl. Acad. Sci. U.S.A. 100 (13), 8007-8012 (2003)
 12808149
 Contact: Tadao Shin-I
 Center For Genetic Resource Information
 National Institute of Genetics
 111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tshin@genes.nig.ac.jp
 A backbone of the vector is basically from pBluescript II (KS), that was in vivo excised from a 1-FLC phage vector (Carninci et al. 2001). 5' end of the cDNA that was digested with XhoI was ligated to SalI site of the vector and the 3' end including polyA tail was ligated to BamHI site of the vector
 vector(5'- ggaagagagagatccacccctgagagattttttttttttttttt-3' was used as a 1st 3' primer, and
 5'-ggttctgagatcgctgctgctcagacagatgactcgagaccggnnnn-3' as 2nd 5'-hairpin primer, giving the following 5' boarder sequence,
 AGGCGCAATCGCGCGAGCTCAATTCGCGAACC). cDNA insert could be amplified with conventional T7 and T3 primers. This full-length

cDNA library was generated according to the method (Nishiyama et al. (2003).
 Protonemata were blended by the POLYTRON, and then the BCDAT medium for 13- 14 days under the continuc These clones are available from RIKEN Bio Resource (http://www.brc.riken.go.jp/lab/epd/Eng/index.html) of Physcomitrella EST clones is available at the PH (http://mos.nibb.ac.jp).
 location/Qualifiers
 1..723
 /organism="Physcomitrella patens subsp. pa:
 /mol_type="mRNA"
 /sub_species="patens"
 /db_xref="taxon:145481"
 /clone="pphn23j11"
 /tissue_type="mixture of chloronemata, cau
 rhizoid-like protonemata"
 /clone_id="normalized full length cDNA li:
 chloronemata, caulonemata and rhizoid-like
 ORIGIN
 Alignment Scores:
 Pred. No.: 3,54e-55 Length: 723
 Score: 676.00 Matches: 126
 Percent Similarity: 83.04% Conservative: 16
 Best Local Similarity: 73.68% Mismatches: 29
 Query Match: 68.70% Indels: 0
 Ds: 3 Gaps: 0
 US-10-716-089-20 (1-188) x BJS98461 (1-723)
 QY 18 ProValMetCysLysAsnLeuCysGlyPhePheGlySerGlnAlaThrM 37
 Db 723 CCTGTGATGTGTAAAGACGTGTGCGGTTCTTTGGACGCCAAGCCATTA 664
 QY 38 SerLysCySyrArGluThrValMetGlnAlaLysMetThrAlaLeu 57
 Db 663 TCGAAGTGTTCAGAGGAGACCGTCATGCAAGCAAGATGACCGCGTG 604
 QY 58 ThrGlnAlaAlaGlnAlaThrSerAlaThrAlaAlaValGlnPro 77
 Db 603 ACTCAAGCGGCTCAGGTCTGCCATCCGCTCATCCAGTGTCACTC 544
 QY 78 HisGluThrLysLeuThrCysGluValAlaGluThrMetIleValPro 97
 Db 543 GAAAGGACAAAGAGCTCTTTGAAGCCGACAGCATGTGATTAACAACGC 484
 QY 98 SerTyrrGlnGlnAspLeuValThrProAlaAlaAlaProGlnAla 117
 Db 483 AGCCACCATCCAGTTGAGGTTGCTCCGGTTACAGTACGAGCCCGAGTTG 424
 QY 118 IleAlaAlaProSerArGProGluProAlaAlaGlySerCysAlaGly 137
 Db 423 GTTGCGACACCCCTCCAGACCGGACCAACCGGTGGATCTTGACAGA 364
 QY 138 LeuThrGlyPheLysCysArGlyAsnLeuTyrrCysAlaLeuHisA 157
 Db 363 TTGACAGGATTCAGTGTCTGTGGGACACTTTTGTGCTTTTACATC 304
 QY 158 LysHisIleThrCysThrTyrrAspTyrrLysAlaAlaGlnGlnAlaAla 177
 Db 303 AAACATAGTTCACATTATGATTAACAAGCTCTGGACAGAGGCTATTG 244
 QY 178 ProLeuValAlaAlaGluLysValAlaLysPhe 188
 Db 243 CTTTATAGTTGTGTGCCGAGAAAGTTGTTAAGTTC 211
 RESULT 14
 BJS86048/c
 LOCUS BJS86048 736 bp mRNA linear 2003
 DEFINITION BJS86048 normalized full length cDNA library, chloronemata and malformed buds Physcomitrella patens cDNA clone pphb29d18 3', mRNA sequence.

ACCESSION BJ586048
 VERSION BJ586048.1 GI:37828036
 KEYWORDS
 SOURCE Physcomitrella patens subsp. patens
 ORGANISM Physcomitrella patens subsp. patens
 Bacteria; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
 Bryopsida; Funariaceae; Funariales; Funariaceae; Physcomitrella.
 REFERENCE 1 (bases 1 to 736)
 AUTHORS Nishiyama, T., Fujita, T., Shin-i, T., Seki, M., Nishide, H.,
 Uchiyama, I., Kamiya, A., Carninci, P., Hayashizaki, Y., Shimozaki, K.,
 Kohara, Y. and Hasebe, M.
 Comparative genomics of Physcomitrella patens gametophytic
 transcriptome and Arabidopsis thaliana: implication for land plant
 evolution
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 100 (13), 8007-8012 (2003)
 PUBMED 12808149
 COMMENT Contact: Tadao Shin-i
 Center For Genetic Resource Information
 National Institute of Genetics
 1111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tshin@genes.nig.ac.jp
 A backbone of the vector is basically from pBluescript II (KS),
 that was in vivo excised from a 1-F1C phage vector (Carninci et al.
 2001). 5' end of the cDNA that was digested with XhoI was ligated
 to SalI site of the vector and the 3' end including polyA tail was
 ligated to BamHI site of the
 vector (5'-gagagagagagatccacacccctgagagagttttttttttttttt-3' was
 used as a 1st 3' primer, and
 5'-ggttctgagtcagtcgtgtccagacgcatgactgagagaccggnnnn-3' as 2nd
 5'-hairpin primer, giving the following 5' boarder sequence,
 AGGCCAATGCGCCGACTGCAATTCGTCGAGACCG). cDNA insert could be
 amplified with conventional T7 and T3 primers. This full-length
 cDNA library was generated according to the method described in
 Nishiyama et al. (2003).
 Protonemata were blended by the POLYTRON, and then cultivated on
 the BCDATG medium for 13-14 days under the continuous light.
 These clones are available from RIKEN Bio Resource Center
 (http://www.brc.riken.go.jp/lab/epd/Eng/index.html). The database
 of Physcomitrella EST clones is available at the PHYSCOBASE
 (http://mes.nihb.ac.jp).
 FEATURES
 source location/Qualifiers
 1..736
 /organism="Physcomitrella patens subsp. patens"
 /mol_type="mRNA"
 /sub_species="patens"
 /db_xref="taxon:145481"
 /clone="pphb29418"
 /ribose_type="mixture of chloronemata, caulonemata and
 malformed buds"
 /clone_1fb="normalized full length cDNA library,
 chloronemata, caulonemata and malformed buds"
 ORIGIN
 Alignment Scores:
 Pred. No.: 2.17e-54 Length: 736
 Score: 668.00 Matches: 125
 Percent Similarity: 82.35% Conservative: 15
 Best Local Similarity: 73.53% Mismatches: 30
 Query Match: 67.89% Indels: 0
 DB: 3 Gaps: 0
 US-10-716-089-20 (1-188) x BJ586048 (1-736)
 Oy 19 ValMetCybLYbAenLeuCyagIyPhePheGlyserGlnAlaThrMerGlnYleuCyaser 38
 Db 734 GTGATGTGTTAGACGCTGTGCGGTTCTTTGCGACGCAAGCAGCTATGGGCTCTGCTCG 675
 Oy 39 LybCybYrYrAGTGTThrValMetGlnAlaLybMetThrAlaLeuAGlGlnAlaThr 58
 Db 674 AAGTGTTCAGGAGACCGTCTATGCAAGCAAGATGACCGCGGTGCGGAGAGAGCCACT 615

Oy 59 GlnAlaAlaGlnAlaThrsAlaThrAlaAlaAlaValAlaGlnProProA 78
 Db 614 CAAGCCGCTCAGAGTGTGCTGCCATCCGTCATCCGGTCTCAGCTCCAG 555
 Oy 79 GluThrLYbLeuThrCyagIyValGluAlaArgThrMetIleValProHisG 98
 Db 554 GAGGACAGAGCTCATTTAGAGCCGACAGCATGTTGATTCMAACCGCTC 495
 Oy 99 TyrgInGlnAspLeuValThrProAlaAlaAlaAlaProGlnAlaVal 118
 Db 494 CACCATCCAGTTGAGAGTTGCTCCGGTTACAGTAGAGCCCCCAGGTTGTA 435
 Oy 119 AlaAlaProSerThrArgProGluProAlaAlaArgCybGlySerCybArgLyb 138
 Db 434 GCCACACCTCCAGACCCGACCAACCCGTCGCGATCTTGCAGAGAG 375
 Oy 139 ThrGlyPheLYbCybGlyserGlybAenLeuYrCybAlaLeuHisArg 158
 Db 374 AGCGATTTCACTGCTGTTGGGACCTTTTGTGTCTTACATCGG 315
 Oy 159 HisThrCybThrTyRAspTYrLYbAlaAlaGlyGlnGlnAlaAlaAla 178
 Db 314 CATAGTTGACCTTATGATTACAAAGCTGCTGAGACAGAGGCTATTGCA 255
 Oy 179 LeuValValAlaGlnLYbValYalYalysPhe 188
 Db 254 TTAGTTGTGGCCGAGAGAGGTTGTTAAGTTCC 225
 RESULT 15
 Bu611574/c 723 bp mRNA linear
 DEFINITION Bu611574 normalized full length cDNA library, chloronemata and rhizoid-like protonemata Physcomitrella subsp. patens cDNA clone phn15g15 3', mRNA sequenc
 Bu611574
 ACCESSION BJ611574.1 GI:37853566
 VERSION BJ611574.1
 KEYWORDS EST.
 SOURCE Physcomitrella patens subsp. patens
 ORGANISM Physcomitrella patens subsp. patens
 Bacteria; Viridiplantae; Streptophyta; Embryophyta; Bryopsida; Funariaceae; Funariales; Funariaceae; Pny
 1 (bases 1 to 723)
 REFERENCE Nishiyama, T., Fujita, T., Shin-i, T., Seki, M., Nishide
 Uchiyama, I., Kamiya, A., Carninci, P., Hayashizaki, Y.,
 Kohara, Y. and Hasebe, M.
 Comparative genomics of Physcomitrella patens gametophytic
 transcriptome and Arabidopsis thaliana: implication for
 evolution
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 100 (13), 8007-8012 (2003)
 PUBMED 12808149
 COMMENT Contact: Tadao Shin-i
 Center For Genetic Resource Information
 National Institute of Genetics
 1111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tshin@genes.nig.ac.jp
 A backbone of the vector is basically from pBluescript II (KS),
 that was in vivo excised from a 1-F1C phage vector
 (Carninci et al. 2001). 5' end of the cDNA that was digested with XhoI
 was ligated to SalI site of the vector and the 3' end including
 polyA tail was ligated to BamHI site of the
 vector (5'-gagagagagagatccacacccctgagagagttttttttttttttt-3' was
 used as a 1st 3' primer, and
 5'-ggttctgagtcagtcgtgtgtccagacgcatgactgagagaccggnnnn-3' as 2nd
 5'-hairpin primer, giving the following 5' boarder sequence,
 AGGCCAATGCGCCGACTGCAATTCGTCGAGACCG). cDNA insert could be
 amplified with conventional T7 and T3 primers. This
 cDNA library was generated according to the method
 described in Nishiyama et al. (2003).
 Protonemata were blended by the POLYTRON, and then
 cultivated on the BCDATG medium for 13-14 days under the continuous
 light. These clones are available from RIKEN Bio Resource

(<http://www.brc.riken.go.jp/lab/epd/Eng/index.html>). The database of Physcomitrella EST clones is available at the PHYSCODbase (<http://mos.nibb.ac.jp/>).

FEATURES

Source

Location/Qualifiers

1..723

/organism="Physcomitrella patens subsp. patens"

/mol_type="mRNA"

/sub_species="patens"

/db_xref="taxon:145481"

/clone="pbn15g15"

/release_type="mixture of chloronemata, caulonemata and rhizoid-like protonemata"

/clone_lib="normalized full length cDNA library", chloronemata, caulonemata and rhizoid-like protonemata"

ORIGIN

Alignment Scores:

Pred. No.:	5,616-52	Length:	723
Score:	643.00	Matches:	125
Percent Similarity:	81.98%	Conservative:	16
Best Local Similarity:	72.67%	Mismatches:	29
Query Match:	65.35%	Indels:	2
DB:	3	Gaps:	0

US-10-716-089-20 (1-188) x BJ611574 (1-723)

```
OY 19 ValMetCyblybAanLeuCybGlypHebGlyserGlnAlaThrMetGlyLeuCybSer 38
   |||||
Db 721 GTGATGTAAAGACGTGCGGTTCTTTGGACCAAGCCACTATGGGTTCTGCTCG 662
   |||||

OY 39 LybCybTYrArGluThrValMetGlnAlaLybMetThrAlaLeuAla-GluGlnAlaTh 58
   |||||
Db 661 AAGTGTTCAGGGAGACCGTCTATGCAGAGAGATGACCGCGGTGGCCGAGCAGCCAC 602
   |||||

OY 58 rGlnAla-AlaGlnAlaThrSerAlaThrAlaAlaAlaValGlnProProAlaProValH 78
   |||||
Db 601 TCAAGCCNGCTCAGCTGCTGCATCCGCTCCAGTCCAGCTCCAGTCTTATGG 542
   |||||

OY 78 l6GluThrlybLeuThrCybGlybValGluArgThrMetLeValProHl6GlnSerSers 98
   |||||
Db 541 AAGAGACAAAGGCTCATTTGAGCCGACAGCATGTGATACACCGCCTCAATCTTCCA 482
   |||||

OY 98 eTYrGlnGlnAspleuValThrProAlaAlaAlaAlaProGlnAlaValLybSerSeri 118
   |||||
Db 481 GCCACCATCCAGTTAGGTGCTCCGGTTACAGTAGCCCCCAGGTGTGATGTCGCCCG 422
   |||||

OY 118 leAlaAlaProSerArGProGluProAsnArgCybGlySerCybArGlybArGValGlyL 138
   |||||
Db 421 TTGCCACACCCCTCCAGACCGGCAACCGGTGGGATCTTGCAGAAAGCGGTGGGAT 362
   |||||

OY 138 eUThrGlyPhelybCybArGCybGlybAanLeuTy-CyAlaLeuHl6ArgTySerAePL 158
   |||||
Db 361 TGACAGGATTCAGTGTGCTGTGGGACCTTTTGTGCTTACATCGATTTCTGACA 302
   |||||

OY 158 ybHl6ThrCybThrTyArbTYrLybAlaAlaGlyGlnGlnAlaAlaAlaLybAlaAsp 178
   |||||
Db 301 AACATAGTTGACCTATGATTTACAAAGGCTGCTGACAGGAGGCTATTCGAAAGCTAATC 242
   |||||

OY 178 rOLeuValValAlaGluLybValValLybPhe 188
   |||||
Db 241 CTTAGTTGTGGCCGAGAAAGTTGTTAAGTTTC 210
   |||||
```

Search completed: December 8, 2005, 18:41:31
Job time : 3756 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 8, 2005, 16:36:03 ; Search time 165 Seconds

(without alignment)
2025.343 Million cell updates/sec

Title: US-10-716-089-20

Perfect score: 984
Sequence: 1 MATEVSESTTSQABEGPVM.....GGEALAKANPLVAKVKVF 188

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlp
-O=/cg2n2_1/USPTO_apool.p/US10716089/rn1 05122005 094344 13172/app query.fasta_1.327
-DB=Issued Patents NA -OPMT=fastext -SUPPLX=ap2n.rn1 -MINMATCH=0.1 -LOOPC=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptio -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10716089 @CGN_1_1_290 @rn1 05122005 094344 13172 -NCPU=6 -ICPU=3
-NO_MMAP -LARGESUBJECT -NEG_SCORES=0 -WAIT -DSPLLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:*

1:	/cg2n2_6/ptodata/1/lna/1_COMB.seq:*
2:	/cg2n2_6/ptodata/1/lna/5_COMB.seq:*
3:	/cg2n2_6/ptodata/1/lna/6A_COMB.seq:*
4:	/cg2n2_6/ptodata/1/lna/6B_COMB.seq:*
5:	/cg2n2_6/ptodata/1/lna/H_COMB.seq:*
6:	/cg2n2_6/ptodata/1/lna/PCTUS_COMB.seq:*
7:	/cg2n2_6/ptodata/1/lna/PP_COMB.seq:*
8:	/cg2n2_6/ptodata/1/lna/RE_COMB.seq:*
9:	/cg2n2_6/ptodata/1/lna/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	984	100.0	804	3	US-09-828-303-12
2	957	97.3	1531	3	US-09-828-303-4
3	367	37.3	2515	3	US-09-244-805-10
4	365	37.1	2425	3	US-09-949-016-810
5	365	37.1	2684	3	US-09-949-016-3693
6	344.5	35.0	1197	2	US-08-861-269-4
7	344.5	35.0	1197	2	US-09-134-596-4
8	344.5	35.0	1197	3	US-09-293-273-4
9	277	28.2	533	3	US-09-640-211A-1804

10	265	26.9	283	3	US-09-313-294A-7446	Seq
11	261	26.5	281	3	US-09-313-294A-1951	Seq
12	219	22.3	211	3	US-09-313-294A-3856	Seq
13	207	21.0	348	3	US-09-640-211A-1742	Seq
14	193.5	19.7	1281	3	US-09-620-312D-849	Seq
15	193	19.6	14524	3	US-09-949-016-12552	Seq
16	193	19.6	14785	3	US-09-949-016-15435	Seq
17	187	19.0	2089	3	US-10-104-047-106	Seq
18	160	16.3	329	3	US-09-640-211A-1702	Seq
19	159.5	16.2	181	3	US-09-313-294A-5600	Seq
20	159	16.2	306	3	US-09-513-999C-9489	Seq
21	148.5	15.1	291	3	US-09-313-294A-6731	Seq
22	148.5	15.1	307	3	US-09-313-294A-6505	Seq
23	148.5	14.9	362	3	US-09-640-211A-1637	Seq
24	146.5	14.9	389	3	US-09-640-211A-1587	Seq
25	146.5	14.7	467	3	US-09-640-211A-1667	Seq
26	145	14.6	563	3	US-09-640-211A-1662	Seq
27	143.5	14.5	315	3	US-09-640-211A-1346	Seq
28	143	14.5	322	3	US-09-640-211A-1337	Seq
29	143	14.5	368	3	US-09-640-211A-1379	Seq
30	142	14.4	313	3	US-09-640-211A-1350	Seq
31	142	14.4	317	3	US-09-640-211A-1723	Seq
32	142	14.4	355	3	US-09-640-211A-1657	Seq
33	142	14.4	406	3	US-09-640-211A-1709	Seq
34	142	14.4	265	3	US-09-313-294A-3339	Seq
35	137.5	14.0	473	3	US-09-640-211A-1740	Seq
36	134	13.6	320	3	US-09-640-211A-1371	Seq
37	129	13.1	330	3	US-09-640-211A-1290	Seq
38	120.5	12.2	290	3	US-09-313-294A-5271	Seq
39	118	11.8	259	3	US-09-313-294A-161	Seq
40	117.5	11.9	1863	3	US-09-902-540-4947	Seq
41	110	11.2	30780	3	US-09-902-540-1243	Seq
42	110	11.2	601	3	US-09-949-016-33790	Seq
43	105	10.7	601	3	US-09-949-016-33790	Seq
44	105	10.7	601	3	US-09-949-016-33790	Seq
45	103.5	10.5	2664	2	US-08-942-819-1	Seq

ALIGNMENTS

RESULT 1
US-09-828-303-12
Sequence 12, Application US/09828303
Patent No. 6677504
GENERAL INFORMATION:
APPLICANT: COSTA E SILVA, OSWALDO DA
APPLICANT: BOHNER, HANS J.
APPLICANT: VAN THIELEN, NOCHA
APPLICANT: CHEN, RUTING
TITLE OF INVENTION: TRANSCRIPTION FACTOR STRESS-RELATED PROT
TITLE OF INVENTION: METHODS OF USE IN PLANTS
FILE REFERENCE: 16313-0030
CURRENT APPLICATION NUMBER: US/09/828,303
CURRENT FILING DATE: 2001-08-20
PRIOR APPLICATION NUMBER: 60/196,001
PRIOR FILING DATE: 2000-04-07
NUMBER OF SEQ ID NOS: 79
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 12
LENGTH: 804
TYPE: DNA
ORGANISM: Phycocytella patens
US-09-828-303-12
Alignment Scores:
Pred. No.: 6,21e-87
Score: 984.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%
DB: 3
Gaps: 0
US-10-716-089-20 (1-188) x US-09-828-303-12 (1-804)

OY 1 MetAlaThrGluArgValSerGlnGluThrThrSerGlnAlaProGluGlyProValMet 20
DB 134 ATGGCCACCGAGCGTGTCTTCAGAGACGACCTTCGAGGCCCTCGAGGCTCCATTATG 193
OY 21 CysAlaSerLeuCysGlyPhePheGlySerGlnAlaThrMetGlyLeuCysSerLeuCys 40
DB 194 TGCAGAAACCTTTGGGGCTTCTTCGGCAGCCAAAGCTACCATGGGGTTGTGCTCGAAGTGC 253
OY 41 TyrArgGluThrValMetGlnAlaLysMetThrAlaLeuAlaGluGlnAlaThrGlnAla 60
DB 254 TACCCAGAGACAGTCATGCAACGAGATGACGGCTTTAGCTGACCAAGCACTCAGGCT 313
OY 61 AlaGlnAlaThrSerAlaThrAlaAlaValGlnProProAlaProValHisGluThr 80
DB 314 GCTCAGGACATCTGCCACAGCTGCTGCTTTCAGCCCCCGCTCTCTGATCATAGACC 373
OY 81 LysLeuThrCysGluValGluArgThrMetCileValProHisGlnSerSerSerTyrln 100
DB 374 AAGCTCACATGGCAGGTTGAGAGAACATGATTTGCCGCATCATCTTCCAGCTATCAA 433
OY 101 GlnApleuValThrProAlaAlaAlaAlaProGlnAlaValLysSerSerTleAlaAla 120
DB 434 CAAGCCTGCTTACCCCGCTGACAGCTGCCCTCAGGAGTGAAGTCTCTATGCGACCT 493
OY 121 ProSerArgProGluProAsnArgCysGlySerCysArgLysArgValGlyLeuThrGly 140
DB 494 CCTCTAGACCCGAGCCCAATGATCGGATCTTTCAGAGACCGTGTGGATTCAGAGAA 553
OY 141 PheLysCysArgCysGlyAsnLeuTyrcysAlaLeuHisArgTyrcysSerAspLysHisThr 160
DB 554 TTTAAGTGCTGCTGTGGCAACCTCTACGCTTACATCGGTACTCGGACAAACACACT 613
OY 161 CysThrTyraPpTyrlsValAlaAlaGlyGlnGluAlaIleAlaLysAlaAspProleuVal 180
DB 614 TGCACATATGACTACAAAGCCGACGAGGAGATTCGAAAGCTAATCTCTGTGC 673
OY 181 ValAlaGluLysValValLysPhe 188
DB 674 GTGGCCGAGAGGTTGTCAAGTTT 697

RESULT 2
US-09-828-303-4
; Sequence 4, Application US/09828303
; Patent No. 6677504
; GENERAL INFORMATION:
; APPLICANT: COSTA E SILVA, OSWALDO DA
; APPLICANT: BOHNER, HANS J.
; APPLICANT: VAN THIELEN, NICHIA
; APPLICANT: CHEN, ROUYING
; TITLE OF INVENTION: TRANSCRIPTION FACTOR STRESS-RELATED PROTEINS AND
; TITLE OF INVENTION: METHODS OF USE IN PLANTS
; FILE REFERENCE: 16313-0030
; CURRENT APPLICATION NUMBER: US/09/828.303
; CURRENT FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: 60/196,001
; PRIOR FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: Patencin Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1531
; TYPE: DNA
; ORGANISM: Physcomitrella patens
US-09-828-303-4

Alignment Scores:
Pred. No.: 6.4e-84 Length: 1531
Score: 957.00 Matches: 187
Percent Similarity: 99.47% Conservative: 0
Best Local Similarity: 99.47% Mismatches: 1
Query Match: 97.26% Indels: 1
DB: 3 Gaps: 0

US-10-716-089-20 (1-188) x US-09-828-303-4 (1-1531)
OY 1 MetAlaThrGluArgValSerGlnGluThrThrSerGlnAlaProGluGly 20
DB 655 ATGGCCACCGAGCGTGTCTTCAGAGACGACCTTCGAGGCCCTCGAGGCTCCATTATG 714
OY 21 CysAlaSerLeuCysGlyPhePheGlySerGlnAlaThrMetGlyLeuCys 40
DB 715 TGCAGAAACCTTTGGGGCTTCTTCGGCAGCCAAAGCTACCATGGGGTTGT 774
OY 41 TyrArgGluThrValMetGlnAlaLysMetThrAlaLeuAlaGluGlnAla 60
DB 775 TACCCAGAGACAGTCATGCAACGAGATGACGGCTTTAGCTGACCAAG 833
OY 61 AlaGlnAlaThrSerAlaThrAlaAlaValGlnProProAlaProVal 80
DB 834 GCTCAGGACATCTGCCACAGCTGCTGCTTTCAGCCCCCGCTCTCTGATCATAGACC 893
OY 81 LysLeuThrCysGluValGluArgThrMetCileValProHisGlnSerSerTyrln 100
DB 894 AAGCTCACATGGCAGGTTGAGAGAACATGATTTGCCGCATCATCTTCCAGCTATCAA 953
OY 101 GlnApleuValThrProAlaAlaAlaAlaProGlnAlaValLysSerSerTleAlaAla 120
DB 954 CAAGCCTGCTTACCCCGCTGACAGCTGCCCTCAGGAGTGAAGTCTCT 1013
OY 121 ProSerArgProGluProAsnArgCysGlySerCysArgLysArgValGlyLeuThrGly 140
DB 1014 CCTCTAGACCCGAGCCCAATGATCGGATCTTTCAGAGACCGTGTGGATTCAGAGAA 1073
OY 141 PheLysCysArgCysGlyAsnLeuTyrcysAlaLeuHisArgTyrcysSerAspLysHisThr 160
DB 1074 TTTAAGTGCTGCTGTGGCAACCTCTACGCTTACATCGGTACTCGGACAAACACACT 1133
OY 161 CysThrTyraPpTyrlsValAlaAlaGlyGlnGluAlaIleAlaLysAlaAspProleuVal 180
DB 1134 TGCACATATGACTACAAAGCCGACGAGGAGATTCGAAAGCTAATCTCTGTGC 1193
OY 181 ValAlaGluLysValValLysPhe 188
DB 1194 GTGGCCGAGAGGTTGTCAAGTTT 1217

RESULT 3
US-09-244-805-10
; Sequence 10, Application US/09244805
; Patent No. 669660
; GENERAL INFORMATION:
; APPLICANT: Worley, Paul F.
; APPLICANT: Lanhahan, Anthony
; APPLICANT: Goetz, Bernard
; APPLICANT: Heilmisch, Holger
; APPLICANT: Kuner, Robin
; APPLICANT: Schiek, Sigrid
; APPLICANT: Nikolich, Karoly
; APPLICANT: Zhukovskii, Eugene
; TITLE OF INVENTION: IMMEDIATE EARLY GENES AND METHODS OF USE
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 10496/004001
; CURRENT APPLICATION NUMBER: US/09/244,805
; CURRENT FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: 60/074,518
; PRIOR FILING DATE: 1998-02-12
; PRIOR APPLICATION NUMBER: 60/074,135
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 2515
; TYPE: DNA
; ORGANISM: Eukaryote
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (414)...(1055)

	:	NUMBER OF SEQ ID NOS:	207012		
	:	SOFTWARE:	Fasteqseq for Windows Version 4.0		
	:	SEQ ID NO	810		
	:	LENGTH:	2425		
	:	TYPE:	DNA		
	:	ORGANISM:	Human		
	:	US-09-949-016-810			
	:	Alignment Scores:			
	:	Pred. No.:	6,48e-26	Length:	2425
	:	Score:	365.00	Matches:	81
	:	Percent Similarity:	52.80%	Conservative:	32
	:	Best Local Similarity:	37.85%	Mismatches:	67
	:	Query Match:	37.09%	Indels:	34
	:	Gaps:	3	Indels:	7
	:	US-10-716-089-20 (1-188) x US-09-949-016-810 (1-2425)			
OY		4 GlnArgValSerGInglUthrThrSerGlnAlaProGluGlyProValM			23
		:::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::			
Db		283 AAAAATATGGCTCAGGAG--ACTAACCAAGACCCTCG--GGGCCCATGC			336
OY		24 LeuCyseGlYPhePheGlyserGlnAlaThMeGlyLeucyyserysc			43
		:::: :::: :::: :::: :::: :::: :::: :::: ::::			
Db		337 GGATGTGGCTTTTATGGAAATCCTAGACCAATGAATTGTTTCAGTTT			396
OY		44 ThrValMetGlnAlaIaymEctHraLaLevalaglUglnAlaThrGlnV			63
		:::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::			
Db		397 CATCTTCAGAGCGACGACAATAATAGTGACAGAATAGCCCAATGGGACAC			456
OY		64 ThrsrAlaThr-----AlalaalaaValGln-----			72
		::: :::: :::: :::: :::: :::: :::: ::::			
Db		457 AACAGTCTACTCTCAGATTCTGCATCTGTACAGAGACAGACACTAGCT			516
OY		73 -----ProProVa			76
		::: :::: :::: :::: :::: :::: :::: ::::			
Db		517 GAAGGTGCTGCTGGCACGACATCTGAAAAATCAAAGAAATGTGCTGTGG			576
OY		77 ValHisglUthrLySleuthrcysgluValgluaGrThrMetileValP			96
		:::: :::: :::: :::: :::: :::: :::: :::: :::: ::::			
Db		577 GTTAATCAGCAATGACAGAAAATGAGACATTTCCAAGAGAGACAAATTA			636
OY		97 SerSetTyrgInlNapleuValThr-----ProAlaalaa			113
		:::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::			
Db		637 ACAGAGGTGTCAAGCGACAGTTGTCACTACGCCCAATCATCATGTTCTC			696
OY		114 VallyserserleialaiaProserArpProgluPro-----AsnA			131
		:::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::			
Db		697 TCTCAGGTGMAAGAAAAAGCTCTGTGAATGGCCCCAACCAAGAAAAACA			756
OY		132 CyARGLYsarValglyleuthrgrlyphelycyebargcysgluysnl			151
		:::: :::: :::: :::: :::: :::: :::: :::: :::: ::::			
Db		757 TGCAAGAAAGAAATTGGTCTTACAGGGTTTGACTGCCGATGTGAANA			816
OY		152 leuHIaGrtySerasplyshiarthrcysthrTYraspyrlysaiaa			171
		:::: :::: :::: :::: :::: :::: :::: :::: :::: ::::			
Db		817 CTTCAACCGTACTCTGCAAGCAACCACTGTCGGTATGATTACAAAGAC			876
OY		172 AlaIIeAlalyAlaenProleuValvaIAlaIGluLysVal 185			
		:::: :::: :::: :::: :::: :::: :::: :::: :::: ::::			
Db		877 AAATCAGAAAAGAGATCCAGATTGTGTGGCTGAAAAAATT 918			
		:::: :::: :::: :::: :::: :::: :::: :::: :::: ::::			
		RESULT 5			
		US-09-949-016-3693			
		/ Sequence 3693, Application US/09949016			
		/ Patent No. 6812339			
		/ GENERAL INFORMATION:			
		/ APPLICANT: VENTER, J. Craig et al.			
		/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED			
		/ WITH HUMAN DISEASE, METHODS OF DETECTIC			
		/ FILE REFERENCE: CL001307			
		/ CURRENT APPLICATION NUMBER: US/09/949, 016			
		/ CURRENT FILING DATE: 2000-04-14			
		THREBE			

[illegible]

Percent Similarity:	51.148	Conservative:	31
Best Local Similarity:	36.998	Mismatches:	67
Query Match:	35.014	Indels:	40
	2	Gaps:	7

US-10-716-089-20 (1-188) x US-09-134-596-4 (1-1197)

QY	2	ALATHRGlu-ArgValSerGlnGluThrThrsSerGlnAlaProGluGly	::
			::
Db	249	GCAATGAGGAAACATGGCTCAAGAAACTATACAGC-----CAAGTAC	::
QY	21	elysAenLeuCySGlyPhePhgGlySerGlnAlaPheMetGlyLeuCyS	::
			::
Db	303	TYCCACTGGCTGTGATTTTATGAGAAACCTCGTACAAATGCAATGTGT	::
QY	41	rARGlu	::
			::
Db	363	TAAAGAACTCTTCAAGACAGAAATGTAATGTAATGTAATGAAATAGCCCA	::
QY	45	lMetGlnAlaLysMetThrAlaLeuAlaGluGlnAlaThr	::
		:::	::
Db	423	TGTCACTAGTCTGTCTGAACTCTTACACAGTTCAATGACAGTCAGCACT	::
QY	60	aAlaGlnAlaThrSerAlaThrAlaAlaAlaValGlnProProAlaPro	::
		:::	::
Db	483	CCAGTCAGCATTTAGACTCTTACATCTTCATCTATGAG---CCACGCTCT	::
QY	80	rLysLeuThrCySGluValGluArgThrMetIleValProHisGlnSer	::
			::
Db	540	GTCACCTTTATACGAA-----TCTGTACATCTTCTCAATTG	::
QY	100	nGlnAProLeuValThrProAlaAlaAlaProGlnAla--ValLys	::
			::
Db	588	TGTGACCAAGCAGTACTGTAAGAACAGAAAGATGTGACAGCTTCAGATCA	::
QY	119	aAlaProSer-----ArgProGlu	::
			::
Db	648	GCAGCATCTGGAAGACAAAGCAAGTCTTTGAAAAACCGAAACAAAAA	::
QY	129	gGlySerCySArgLysArgValGlyLeuThrGlyPheLysCySArgCyS	::
		:::	::
Db	708	TTTCACTGTGACAGGAAAGAGTGGACTTACGTGGTTTGAATCCCGGT	::
QY	149	rCySAlaLeuHisArgTySerAspLysHisThrCySArgThrTyAspTy	::
		:::	::
Db	768	CTGTGTGTGACACCGTTACTCAGATGTACACAAATGTCTTACCAATTAC	::
QY	169	yGlnGluAlaIleAlaLysAlaAsnProLeuValValaGluLysVal	::
			::
Db	828	TGCTGAGAAAATCAGAAAAGAAAATCCAGTAGTGTGTGTAAGAAAGATC	::

RESULT 8
US-09-293-273-4
Sequence 4, Application US/09293273
Patent No. 6057112
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Corley, Neil C.
APPLICANT: Lal, Preeti
APPLICANT: Shah, Purni
TITLE OF INVENTION: UBIQUITIN CONJUGATION PROTEINS
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/293,273
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/134,596
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0302 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1197 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BRAINOT09
CLONE: 2151473
US-09-293-273-4

Alignment Scores:

Pred. No.:	2,486-24	Length:	1197
Score:	344.50	Matches:	81
Percent Similarity:	51.14%	Conservative:	31
Best Local Similarity:	36.99%	Mismatches:	67
Query Match:	35.01%	Indels:	40
DB:	3	Gaps:	7

US-10-716-089-20 (1-188) x US-09-293-273-4 (1-1197)

```
QY 2 AAlarhrglu-ArgValSerGlnGluThrThrsSerGlnAlaProGlnGlyProValMetCy 21
DB 249 GCAACTGAGGAACATGGGCTCAAGAACTAATCAAGC-----CAAGGCGCTATGCTTTG 302
QY 21 gLyAsnleuCygGlyPhepneGlySerGlnAlaThrMetGlyLeuCySerLyCyety 41
DB 303 TTCCA CTGGCTGTGGATTTTATGAAACCTCGTCAAAATGGCATGTGTTCAGTATGCTA 362
QY 41 rArgGlu-----ThrVa 45
DB 363 TAAAGAACTCTTCAAGACAGATATGTAATGTATGAATTAACCCACTGCACCTC 422
QY 45 lmerGlnAlaLyMetThrAlaLeuAlaGluGlnAlaThr-----GlnAl 60
DB 423 TGTCAGTAGTCTGTGATCTTTTACAGTTCAATGCACAGATGGCAGTGTCCAGAAAGC 482
QY 60 aAlaGlnAlaThrSerAlaThrAlaAlaAlaValGlnProProAlaProValHisGluTh 80
DB 483 CCAGTCACACATTTAGACTTCACTCTTCATCTATGAG--CCACACCTGTATCAAAATCA 539
QY 80 rLySerLeuThrCyGluValAlaGluArgThrMetIleValProHisGlnSerSerSerTyG 100
DB 540 GTCACTTTTATAGAA-----TCTGTACATCTTCTCAATTGGACAGTACATC 587
QY 100 nglnAspleuValThrProAlaAlaAlaAlaProGlnAla---ValLySerSerIleAl 119
DB 588 TGTGACAAAGACAGTACCTGAAGAAAGAGATGTGACAGGCTTCAGATACACACAGACA 647
QY 119 aAlaProSer-----ArgProGluProSerArgCy 129
DB 648 GCAGCACACTGTAAGACAAAGCAAGTCTTGTAAAAACCGAAACAAAAAGAAATCGCTG 707
QY 129 gLySerCySaArgLyAsrValGlyLeuThrGlyPheLyCySaArgCyGlyAsnLeuTy 149
DB 708 TTTCACTGTGCAAGAAAGAGTGGACTTACTGGGTTTGAATCCCGGTGTGAAATGTTTA 767
QY 149 rCyAlaLeuHisArgTySerAspLySerHisThrCySerThrTyAspTyTyLyAlaAlaG 169
DB 169
```

```
DB 768 CTGTGTGTACACCGCTTACTGATGTACACATTCGCTTACATTTAC
QY 169 yGlnGluAlaIleAlaLyAsnProLeuValAlaGlnLySerVal
DB 828 TGTGAGAAATCAGAAAGAAATCCAGTATGTTGTGTGAAAGATC
```

RESULT 9

US-09-640-211A-1804

Sequence 1804, Application US/09640211A
Patent No. 6833446
GENERAL INFORMATION:
APPLICANT: Wood, Marion
APPLICANT: Shenk, Michael A.
APPLICANT: McGrath, Annette
APPLICANT: Glenn, Matthew
TITLE OF INVENTION: Compositions and Methods for the
TITLE OF INVENTION: Modification of Gene Transcription
FILE REFERENCE: 11000.1021CIU
CURRENT APPLICATION NUMBER: US/09/640,211A
CURRENT FILING DATE: 2000-08-16
NUMBER OF SEQ ID NOS: 2368
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1804
LENGTH: 533
TYPE: DNA
ORGANISM: Pinus radiata
US-09-640-211A-1804

Alignment Scores:

Pred. No.:	3,16e-18	Length:	533
Score:	277.00 <td>Matches:</td> <td>19</td>	Matches:	19
Percent Similarity:	54.25% <td>Conservative:</td> <td>64</td>	Conservative:	64
Best Local Similarity:	41.83% <td>Mismatches:</td> <td>46</td>	Mismatches:	46
Query Match:	28.15% <td>Indels:</td> <td>24</td>	Indels:	24
DB:	3 <td>Gaps:</td> <td>5</td>	Gaps:	5

US-10-716-089-20 (1-188) x US-09-640-211A-1804 (1-533)

```
QY 5 ArgValSerGlnGlu-----ThrThrsrG
DB 114 AGGATGATCAAGAAACTGGAACATCGAGCTGATGCGCTGCTCC
QY 17 gLyProValMetCyLyAsnleuCygGlyPhepneGlySerGlnAlaThr
DB 174 GGGCACTCTTTCGGCAATACATCGCGGCTTTTGGCACTTGGCAAC
QY 37 CySerLyCyTyArgGluThrValMetGlnAlaLyMetThrAlaLe
DB 234 TGTTCGAAATGTTACAGGATCTGATTAAGAAG-----GAGGCCCT
QY 57 AAlarhrglnAlaGlnAlaThrSerAlaThrAlaAlaValGlnP
DB 285 GCAATGGCCGCGTTGAGAGTCAATTGCGCGGCTTCCGATGGAGG
QY 77 ValHisGluThrLySerLeuThrCyGluValGluArgThrMetIleValP
DB 345 CTTTCAAGCAGAGATGTTTGTGTGAAACAAGCCGTGACCGATCTCC
QY 94 HisGlnSerSerSerTyGlnGlnAspleuValThrProAlaAlaAla
DB 405 CAAGCTCTGTCAAGTCACTTGGCTGATATA-----GGTTATCTTCTT
QY 114 ValLySerSerIleAlaAlaProSerArgProGluProAsnArgCyG
DB 456 -----CTCTGCGCAAACTCTTAACCGGTGCT
QY 134 LySaArgValGlyLeuThrGlyPheLyCySaArgCyGly 146
DB 495 AAGCAGTCGCTGACCGCGCTTCAAAATGTGCTGCGGA 533
```

RESULT 10

US-09-313-294A-7446

Sequence 7446, Application US/09313294A

```
/ Patent No. 6476212
/ GENERAL INFORMATION:
/ APPLICANT: Laljudi, Raghunath V.
/ APPLICANT: Ito, Laura Y.
/ APPLICANT: Sherman, Bradley K.
/ TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
/ FILE REFERENCE: PL-0017 US
/ CURRENT APPLICATION NUMBER: US/09/313,294A
/ CURRENT FILING DATE: 1999-05-14
/ NUMBER OF SEQ ID NOS: 7600
/ SOFTWARE: PERL Program
/ SEQ ID NO 7446
/ LENGTH: 283
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ NAME/KEY: misc_feature
/ OTHER INFORMATION: Incyte ID No. 6476212 700381739H1
/ LOCATION: 216
/ NAME/KEY: unsure
/ OTHER INFORMATION: a, t, c, g, or other
/ US-09-313-294A-7446

Alignment Scores:
Pred. No.: 1,988-17 Length: 283
Score: 265.00 Matches: 44
Percent Similarity: 85.48% Conservative: 9
Best Local Similarity: 70.97% Mismatches: 9
Query Match: 26.93% Indels: 0
DB: 3 Gaps: 0

US-10-716-089-20 (1-188) x US-09-313-294A-7446 (1-283)

QY 126 ProaenargCyeglySerCyaaRglYsaRgVaIglYleuThGlyPheLYsCyaaRgCys 145
DB 2 CCAACCGGTGGCGAACCTGTAGGAAGCGTGTGGGTTTAACTGCCGATGC 61
QY 146 GYAaenLeuTYrCYaAlaLeuHIsaRgTYrSerAaRlySHIsThrCYaThTYrAspTYr 165
DB 62 GGGAAcAcTGTACTGTTCGATGACCGCTACTCCGACAAACAGACACTGCAGTTCGACTAT 121
QY 166 LysAlaAlaGlyGlnGlnJuaIaIleAlaLysAlaAenProLeuValValAlaGlnLYsVal 185
DB 122 CGAACTGCAGCTAGGAGCCCTATCCCAAGGCCCAATCCAGTGTGTAAGCGAAGACTT 181
QY 186 ValLYs 187
DB 182 GACAAG 187

RESULT 11
US-09-313-294A-1951
/ Sequence 1951, Application US/09313294A
/ Patent No. 6476212
/ GENERAL INFORMATION:
/ APPLICANT: Laljudi, Raghunath V.
/ APPLICANT: Ito, Laura Y.
/ APPLICANT: Sherman, Bradley K.
/ TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
/ FILE REFERENCE: PL-0017 US
/ CURRENT APPLICATION NUMBER: US/09/313,294A
/ CURRENT FILING DATE: 1999-05-14
/ NUMBER OF SEQ ID NOS: 7600
/ SOFTWARE: PERL Program
/ SEQ ID NO 1951
/ LENGTH: 281
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ NAME/KEY: misc_feature
/ OTHER INFORMATION: Incyte ID No. 6476212 700551796H1
/ US-09-313-294A-1951

Alignment Scores:
```

```
Pred. No.: 4,818-17 Length: 281
Score: 261.00 Matches: 43
Percent Similarity: 87.10% Conservative: 11
Best Local Similarity: 69.35% Mismatches: 8
Query Match: 26.52% Indels: 0
DB: 3 Gaps: 0

US-10-716-089-20 (1-188) x US-09-313-294A-1951 (1-281)

QY 126 ProaenargCyeglySerCyaaRglYsaRgVaIglYleuThGlyPheL 145
DB 10 CCAAGCCGGTGGCGCGCCCTGCAGGAAGCGTGTGGGCTGACCGGATTTT 69
QY 146 GYAaenLeuTYrCYaAlaLeuHIsaRgTYrSerAaRlySHIsThrCYaT 165
DB 70 GGGAAcAcTGTACTGTTCGATGACCGCTACTCCGACAAACAGACTGTC 129
QY 166 LysAlaAlaGlyGlnGlnJuaIaIleAlaLysAlaAenProLeuValVal 185
DB 130 CGAACTGCAGCAAGGAGCGCGATTCGCAAGGCCCAATCCGTGTGAGG 189
QY 186 ValLYs 187
DB 190 GACAAG 195

RESULT 12
US-09-313-294A-3856
/ Sequence 3856, Application US/09313294A
/ Patent No. 6476212
/ GENERAL INFORMATION:
/ APPLICANT: Laljudi, Raghunath V.
/ APPLICANT: Ito, Laura Y.
/ APPLICANT: Sherman, Bradley K.
/ TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVE
/ FILE REFERENCE: PL-0017 US
/ CURRENT APPLICATION NUMBER: US/09/313,294A
/ CURRENT FILING DATE: 1999-05-14
/ NUMBER OF SEQ ID NOS: 7600
/ SOFTWARE: PERL Program
/ SEQ ID NO 3856
/ LENGTH: 211
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ NAME/KEY: misc_feature
/ OTHER INFORMATION: Incyte ID No. 6476212 700282265H2
/ NAME/KEY: unsure
/ LOCATION: 41, 60, 99, 116, 118, 135, 154
/ OTHER INFORMATION: a, t, c, g, or other
/ US-09-313-294A-3856

Alignment Scores:
Pred. No.: 4,066-13 Length: 211
Score: 219.00 Matches: 39
Percent Similarity: 79.37% Conservative: 11
Best Local Similarity: 61.90% Mismatches: 12
Query Match: 22.26% Indels: 1
DB: 3 Gaps: 0

US-10-716-089-20 (1-188) x US-09-313-294A-3856 (1-211)

QY 126 ProaenargCyeglySerCyaaRglYsaRgVaIglYleuThGlyPheL 145
DB 9 CCAAGCCGGTGGCGCGCCCTGCAGGAAGCGTGTGGGCTGACCGGATTTT 68
QY 146 GYAaenLeuTYrCYaAlaLeuHIsaRgTYrSerAaRlySHIsThrCYa 165
DB 69 GGGAAcAcTGTACTGTTCGATGACCGCTACTCCGACAAACAGACTG 128
QY 165 LysAlaAlaGlyGlnGlnJuaIaIleAlaLysAlaAenProLeuValVal 185
DB 129 TCGACNCGCAGCAAGGAGCGCGATTCGCAAGGCCCAATCCGTGTGAGG 188
```

```
OY      185 lvallys 187
      :   |||
Db      189 CGACAG 195

RESULT 13
US-09-640-211A-1742
; Sequence 1742, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Modifications and Methods for the
; FILE REFERENCE: 11000.1021C1U
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: PastSeq for Windows Version 4.0
; SEQ ID NO 1742
; LENGTH: 348
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-640-211A-1742

Alignment Scores:
Pred. No.:      1,192-11      Length:      348
Score:          207.00        Matches:      45
Percent Similarity: 55.45%    Conservative: 11
Best Local Similarity: 44.55%  Mismatches:  21
Query Match:    21.04%       Indels:       24
DB:             3           Gaps:         24

US-10-716-089-20 (1-188) x US-09-640-211A-1742 (1-348)

OY      91 lvalProHleGlnSerSerSerTyrgInGlnAPLeuValThProAlaAlaAla 110
      :   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      60 TTGGTGTCC-----TCAGTAGCGACGCTTCCAAATGTGTGAAGACGTAGCACACAG 110

OY      111 ProGlnAlaValLysSerSerSerTleAlaAProSer----- 122
      :   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      111 GAGACGGCCATTGGACCC-----GCTCCATCCAGCTCCGCAAGACAGAGGTGAG 161

OY      123 -----ArgProGlnProAlaAlaArgCyglLysCys 132
      :   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      162 AAACATGAACAGACGAGGAGGAGACGCTGAGAAAGCAAGAGAGGTGCTGT 221

OY      133 ArgLysArgValGlyLeuThrGlyPheLysCyArgCyglYAsnLeuTyrCyAlaLeu 152
      :   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      222 AACACGGCGGGTGGGCTGACGGGCTTTAAGTCCCTGTGTGACCTTTTCTGCCCTCAG 281

OY      153 HlsArgTyrSerAPLysHlsThrCyThrTyrAPTyTTLysAlaAlaGlyGlnGluAla 172
      :   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      282 CACAGGTACTGTGATATGATGATGATGCTCTTTTGACTACAAAGATCGCGGCGCTGCC 341

OY      173 lIe 173
      :   |||
Db      342 ATT 344

RESULT 14
US-09-620-312D-849
; Sequence 849, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
```

```
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yungting
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhilwei
; APPLICANT: John Tillinphast
; APPLICANT: Drmanac, Radote T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; FILE REFERENCE: Polypeptides
; CURRENT APPLICATION NUMBER: US/09/620,312D
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: PE_FL_genes Version 1.0
; SEQ ID NO 849
; LENGTH: 1281
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (416)..(1033)
US-09-620-312D-849

Alignment Scores:
Pred. No.:      1,45e-09      Length:      1281
Score:          193.50        Matches:      50
Percent Similarity: 43.09%    Conservative: 31
Best Local Similarity: 26.60%  Mismatches:  72
Query Match:    19.66%       Indels:       35
DB:             3           Gaps:         7

US-10-716-089-20 (1-188) x US-09-620-312D-849 (1-1281)

OY      12 SerGlnAlaProGlnGlyProValMetCyLysAsnLeuCyGlyPheP 31
      :   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      440 AGCAAGGCGCCACGCTGCCCTCGCTGCC-----TCCGGCTTCT 493

OY      32 AlaThrMetGlyLeuCySerLysCyEtyrArgGlu----- 43
      :   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      494 AAGACTATGATCTCTGTCTCAAAATGCTTGTCTGATTTTCAAAAGAAC 553

OY      44 -----ThrValMetGlnAlaLysMetThrAlaLeuAlaGluG 59
      :   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      554 GATTCGGCTCCAGTACAACTAACGACGCAATCAGATTGTTTCCGAAG 613

OY      60 AlaAlaGlnAlaThrSerAlaThrAlaAlaAlaValAlaPro----- 76
      :   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      614 GACAACAACAATACCTGATTAACCAAGCCAACTCTTACTGATCCAGCCAGC 673

OY      77 ValHlsGluThrLysLeuThrCyGluValGluArgThrMetLysValP 96
      :   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      674 -----ACAGAACTGAATGTAATCTGC 706

OY      97 SerSerTyrgInGlnAPLeuValThProAlaAlaAlaAProGlnA 116
      :   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      707 GAGTATTCACAGTCTGAGAAATGAGGCTTCACACAGTAAACCGCACGAC 766

OY      117 Ser-----lAlaAlaAlaProSerArgProGlnProAla 131
      :   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      767 ACGGAACGCTCCAGGAACCAAGTCGATTTAAACGAAGAAGTCAGAGTCC 826

OY      132 CysArgLysArgValGlyLeuThrGlyPheLys-----CyArgCy 148
      :   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      827 TGCCAACCAAACTGAGACTGTGTGACAGAGAAATTTGGAGATGTGTGCTC 886

OY      149 TyrCyAlaLeuHlsArgTyrSerAPLysHlsThrCyThrTyrAPTy 168
      :   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
Db      887 TCTGTATGTATCATGCGCTCCCGGAGCAGCAGTGCATTCGACCATGCGCGCT 946
Qy      169 G1Y---G1NG1UA1A11eAla1yS 175
          |||:::|||||
Db      947 GCGCGGAGGAGGCGCATCATGAAA 970

RESULT 15
US-09-949-016-12552
/ Sequence 12552, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: C1001307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ CURRENT FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO: 12552
/ LENGTH: 14524
/ TYPE: DNA
/ ORGANISM: Human
US-09-949-016-12552

Alignment Scores:
Pred. No.:      4.39e-08      Length:      14524
Score:          193.00        Matches:      33
Percent Similarity: 76.60%    Conservative: 3
Best Local Similarity: 70.21%  Mismatches:    11
Query Match:     19.61%      Indels:        0
DB:               3          Gaps:          0

US-10-716-089-20 (1-188) x US-09-949-016-12552 (1-14524)

Qy      139 ThrGlyPheLysCyArgCyAGlyAsnLeuTyrCyAlaLeuH1aArgTyrSerAspLys 158
          |||:::|||||
Db      10877 ACAGGCTTTGACTGCCGAGTGTGAATTGTTTGTGACTTCACCGTTACTCTGACAG 10936
          |||:::|||||
Qy      159 HisThrCysThrTyrAspTyrLysAlaAlaGlyGlnGluAla11eAlaLysAlaAsnPro 178
          |||:::|||||
Db      10937 CACAACGTGTCGATGATTACAAAGCAGAACTGCACGCAAAATCAGAAAGAGATCCA 10996
          |||:::|||||
Qy      179 LeuVal1Val1AlaGluLysVal 185
          |||:::|||||
Db      10997 GTTGTGTGGCTGAAAAAATT 11017
```

Search completed: December 8, 2005, 18:44:31
Job time : 172 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 8, 2005, 16:36:37 ; Search time 791 Seconds
(without alignments)

1965.415 Million cell updates/sec

Title: US-10-716-089-20

Perfect score: 984
Sequence: 1 WATERVSQETTSQABEGPVM.....GGEAIKANPLVAKVKVF 188

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

MODEL=frame+ p2n.model -DEV=xlp
-O=/cg2_1/USPTO.spool.p/US10716089/crunat.05122005.094344.13199/app_query.fasta_1.327
-DB=Published Applications NA.Main -QFMT=faaetap -SUFFIX=p2n.rnpbm
-MINMATCH=0.1 -DOOPCL=0 -DOOPEXT=0 -UNITS=bits -START=1 -END=1
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR_SCORE=spc
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pcio -NORM=ext
-HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10716089 @CGN 1 1549 @runat.05122005.094344.13199 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEOUTERY -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -NARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Published Applications NA.Main:*

1: /cg2_6/prodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cg2_6/prodata/1/pubpna/US08_PUBCOMB.seq:*
3: /cg2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*
4: /cg2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*
5: /cg2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
6: /cg2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
7: /cg2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
8: /cg2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
9: /cg2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
10: /cg2_6/prodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	984	100.0	804	3	US-09-828-303-12
2	984	100.0	804	7	US-10-716-089-12
3	957	97.3	1531	3	US-09-828-303-4
4	957	97.3	1531	7	US-10-716-089-4
5	425.5	43.2	945	7	US-10-425-114-26146
6	425.5	43.2	947	8	US-10-425-115-71431
7	425.5	43.2	1142	7	US-10-767-701-15574

8	419.5	42.6	2671	7	US-10-437-963-21113	Seq
9	416.5	42.3	1032	7	US-10-425-114-27574	Seq
10	416.5	42.3	1171	7	US-10-437-963-37809	Seq
11	416.5	42.3	1249	7	US-10-437-963-37808	Seq
12	415	42.2	736	7	US-10-425-114-22047	Seq
13	415	42.2	776	7	US-10-425-114-28524	Seq
14	415	42.2	821	7	US-10-425-114-28524	Seq
15	415	42.2	826	7	US-10-425-114-21992	Seq
16	415	42.2	841	7	US-10-425-114-25691	Seq
17	415	42.2	846	7	US-10-425-114-25727	Seq
18	415	42.2	853	7	US-10-425-114-28785	Seq
19	415	42.2	863	7	US-10-425-114-27172	Seq
20	415	42.2	863	7	US-10-425-114-28540	Seq
21	415	42.2	868	7	US-10-425-114-15838	Seq
22	415	42.2	868	7	US-10-425-114-28768	Seq
23	415	42.2	872	7	US-10-425-114-26442	Seq
24	415	42.2	872	7	US-10-425-114-27856	Seq
25	415	42.2	873	7	US-10-425-114-27861	Seq
26	415	42.2	876	7	US-10-425-114-27629	Seq
27	415	42.2	886	7	US-10-425-114-6944	Seq
28	415	42.2	894	7	US-10-425-114-23540	Seq
29	415	42.2	898	7	US-10-425-114-32027	Seq
30	415	42.2	899	7	US-10-425-114-25461	Seq
31	415	42.2	902	7	US-10-425-114-27516	Seq
32	415	42.2	904	7	US-10-425-114-21878	Seq
33	415	42.2	918	7	US-10-425-114-32045	Seq
34	415	42.2	920	7	US-10-425-114-22547	Seq
35	415	42.2	924	7	US-10-425-114-31655	Seq
36	415	42.2	926	7	US-10-425-114-27619	Seq
37	415	42.2	928	7	US-10-425-114-27609	Seq
38	415	42.2	931	7	US-10-425-114-28237	Seq
39	415	42.2	932	7	US-10-425-114-25774	Seq
40	415	42.2	942	7	US-10-425-114-28518	Seq
41	415	42.2	943	7	US-10-425-114-28353	Seq
42	415	42.2	948	7	US-10-425-114-5385	Seq
43	415	42.2	953	7	US-10-425-114-27813	Seq
44	415	42.2	956	7	US-10-425-114-27789	Seq
45	415	42.2	959	7	US-10-425-114-22492	Seq

ALIGNMENTS

RESULT 1
US-09-828-303-12
Sequence 12, Application US/09828303
Patent No. US20020102695A1
GENERAL INFORMATION:
APPLICANT: COSTA E SILVA, OSWALDO DA
APPLICANT: BOHNETT, HANS J.
APPLICANT: VAN THIELEN, NOCHA
APPLICANT: CHEN, ROUYING
TITLE OF INVENTION: TRANSCRIPTION FACTOR STRESS-RELATED PROT
FILE REFERENCE: 16313-0030
CURRENT APPLICATION NUMBER: US/09/828,303
CURRENT FILING DATE: 2001-08-20
PRIOR APPLICATION NUMBER: 60/196,001
PRIOR FILING DATE: 2000-04-07
NUMBER OF SEQ ID NOS: 79
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 12
LENGTH: 804
TYPE: DNA
ORGANISM: Physcomitrella patens
US-09-828-303-12

Alignment Scores:
Pred. No.: 4.78e-105
Score: 984.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%
DB: 3
Length: 804
Matches: 188
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0

US-10-716-089-20 (1-188) x US-09-828-303-12 (1-804)

QY 1 MetAaThrGluArgValSerGlnGluThrThrSerGlnAlaProGluGlyProValMet 20
DB 134 ATGGCCACCGACCGTGTCTTCAGAGAGACGACCTCGACGAGCCCTGAGGGGTCCAGTTATG 193
QY 21 CysIlyAsnLeuCyGlyPhePheGlySerGlnAlaThrMetGlyLeuCySerIlySer 40
DB 194 TGCAGAAACCTTTCGGCTTTCGACAGCCAGCTACCATGGGTGTGCTCGAAGTGC 253
QY 41 TyrArgGluThrValMetGlnAlaIlySerMetThrAlaLeuAlaGluGlnAlaThrGlnAla 60
DB 254 TACCGAGAGACGTCATGCAAGCGAAGATGACGGCTTTAGCTGACAGCAAGCCACTCAGGCT 313
QY 61 AlaGlnAlaThrSerAlaThrAlaAlaAlaValGlnProProAlaProValHisGluThr 80
DB 314 GCTCAGGCGACATCTGCCACACCTGCTGTTCAGCCCCCGCTCTGTACATGAGACC 373
QY 81 LysLeuThrCysGluValGluArgThrMetIleValProHisGlnSerSerSerTyrGln 100
DB 374 AAGCTCACATGCGAGGTGTGAGAAACAATGATTGCGCATCAATCTTCCAGCTATCA 433
QY 101 GlnAspLeuValThrProAlaAlaAlaAlaProGlnAlaValIlySerSerIleAlaAla 120
DB 434 CAAGACCTGTTACCCCGCTGACGCTGCCCTCAGGCGCATGAATCCTATCGCAGCT 493
QY 121 ProSerArgProGluProAsnArgCysGlySerCysArgIlyArgValGlyLeuThrGly 140
DB 494 CCTCTAGACCCGAGCCCAATGATGCCGATCTTGACGAGAACGTTGTGATTGACAGCA 553
QY 141 PheIlyCysArgCysGlyAsnLeuTyrCysAlaLeuHisArgTyrSerAspIlySerHisThr 160
DB 554 TTAAAGTGTGCTGTGGCAACCTCTACCTGCGCTTACATCGGTACTCGACAAACACT 613
QY 161 CysThrTyrAspTyrIlyAlaAlaGlyGlnGluAlaIleAlaIlyAlaAsnProLeuVal 180
DB 614 TGCACATATGACTACAAAGCCGACGGCGAGAGCATGGCAAAAGCTATCTCTTGTGC 673
QY 181 ValAlaGluIlyValIlyAspPhe 188
DB 674 GTGGCCGAGAAAGTTGTCAAGTTT 697

RESULT 2

US-10-716-089-12
; Sequence 12, Application US/10716089
; Publication No. US20040107463A1
; GENERAL INFORMATION:
; APPLICANT: COSTA E SILVA, OSWALDO DA
; APPLICANT: BOHNER, HANS J.
; APPLICANT: VAN THIELEN, NOCHA
; APPLICANT: CHEN, ROUYING
; TITLE OF INVENTION: TRANSCRIPTION FACTOR STRESS-RELATED PROTEINS AND
; TITLE OF INVENTION: METHODS OF USE IN PLANTS
; FILE REFERENCE: 16313-0030
; CURRENT APPLICATION NUMBER: US/10/716,089
; CURRENT FILING DATE: 2003-11-18
; PRIOR FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 12
; LENGTH: 804
; TYPE: DNA
; ORGANISM: Physcomitrella patens
US-10-716-089-12

Alignment Scores:
Pred. No.: 4 78e-105 Length: 804
Score: 984.00 Matches: 188
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0

DB: 7 Gaps: 0

US-10-716-089-20 (1-188) x US-10-716-089-12 (1-804)

QY 1 MetAaThrGluArgValSerGlnGluThrThrSerGlnAlaProGluGlyProValMet 20
DB 134 ATGGCCACCGACCGTGTCTTCAGAGAGACGACCTCGACGAGCCCTGAGGGGTCCAGTTATG 193
QY 21 CysIlyAsnLeuCyGlyPhePheGlySerGlnAlaThrMetGlyLeuCySerIlySer 40
DB 194 TGCAGAAACCTTTCGGCTTTCGACAGCCAGCTACCATGGGTGTGCTCGAAGTGC 253
QY 41 TyrArgGluThrValMetGlnAlaIlySerMetThrAlaLeuAlaGluGlnAlaThrGlnAla 60
DB 254 TACCGAGAGACGTCATGCAAGCGAAGATGACGGCTTTAGCTGACAGCAAGCCACTCAGGCT 313
QY 61 AlaGlnAlaThrSerAlaThrAlaAlaAlaValGlnProProAlaProValHisGluThr 80
DB 314 GCTCAGGCGACATCTGCCACACCTGCTGTTCAGCCCCCGCTCTGTACATGAGACC 373
QY 81 LysLeuThrCysGluValGluArgThrMetIleValProHisGlnSerSerSerTyrGln 100
DB 374 AAGCTCACATGCGAGGTGTGAGAAACAATGATTGCGCATCAATCTTCCAGCTATCA 433
QY 101 GlnAspLeuValThrProAlaAlaAlaAlaProGlnAlaValIlySerSerIleAlaAla 120
DB 434 CAAGACCTGTTACCCCGCTGACGCTGCCCTCAGGCGCATGAATCCTATCGCAGCT 493
QY 121 ProSerArgProGluProAsnArgCysGlySerCysArgIlyArgValGlyLeuThrGly 140
DB 494 CCTCTAGACCCGAGCCCAATGATGCCGATCTTGACGAGAACGTTGTGATTGACAGCA 553
QY 141 PheIlyCysArgCysGlyAsnLeuTyrCysAlaLeuHisArgTyrSerAspIlySerHisThr 160
DB 554 TTAAAGTGTGCTGTGGCAACCTCTACCTGCGCTTACATCGGTACTCGACAAACACT 613
QY 161 CysThrTyrAspTyrIlyAlaAlaGlyGlnGluAlaIleAlaIlyAlaAsnProLeuVal 180
DB 614 TGCACATATGACTACAAAGCCGACGGCGAGAGCATGGCAAAAGCTATCTCTTGTGC 673
QY 181 ValAlaGluIlyValIlyAspPhe 188
DB 674 GTGGCCGAGAAAGTTGTCAAGTTT 697

RESULT 3

US-09-828-303-4
; Sequence 4, Application US/09828303
; Patent No. US20020102695A1
; GENERAL INFORMATION:
; APPLICANT: COSTA E SILVA, OSWALDO DA
; APPLICANT: BOHNER, HANS J.
; APPLICANT: VAN THIELEN, NOCHA
; APPLICANT: CHEN, ROUYING
; TITLE OF INVENTION: TRANSCRIPTION FACTOR STRESS-RELATED PROTI
; TITLE OF INVENTION: METHODS OF USE IN PLANTS
; FILE REFERENCE: 16313-0030
; CURRENT APPLICATION NUMBER: US/09/828,303
; CURRENT FILING DATE: 2001-08-20
; PRIOR FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1531
; TYPE: DNA
; ORGANISM: Physcomitrella patens
US-09-828-303-4

Alignment Scores:
Pred. No.: 1.58e-101 Length: 1531
Score: 957.00 Matches: 187
Percent Similarity: 99.47% Conservative: 0
Best Local Similarity: 99.47% Mismatches: 1

Query Match: 97.26% Indels: 1
DB: 3 Gaps: 0
US-10-716-089-20 (1-188) x US-09-828-303-4 (1-1531)

QY 1 MetAlaThrGluArgValSerGlnGluThrThrSerGlnAlaProGlnGluProValMet 20
DB 655 ATGGCCACCGAGCGTGTCTCTCAGAGAGACCTCCGAGGCCCTTAGGGCTCAGTTATG 714
QY 21 CysLysAsnLeuCysGlyPhePheGlySerGlnAlaThrMetGlyLeuCysSerLysCys 40
DB 715 TGCAGAACTTTGCGGCTTCTTCCGACGACGACTACCATGGGGTGTGCTGCAAGTGC 774
QY 41 TyrArgGluThrValMetGlnAlaLysMetThrAlaLeuAlaGluGlnAlaThrGlnAla 60
DB 775 TACCGAGACAGTCATGACGCG-AAGATGACGGCTTTAGCTGAGCAACCTCAGGCT 833
QY 61 AlaGlnAlaThrSerAlaThrAlaAlaValGlnProProAlaProValHisGluThr 80
DB 834 GCTCAGGCGACATCTGCCACAGCTGCTGTTCAAGCCCCGCTCTGTACATGAGACC 893
QY 81 LysLeuThrCysGluValGluArgThrMetIleValProHisGlnSerSerSerTyrGln 100
DB 894 AAGCTCACATGCGAGGTGAGAGAACATGATGTGCGCCATCAATCTTCCAGCTATCA 953
QY 101 GlnAspLeuValThrProAlaAlaAlaProGlnAlaValLysSerSerIleAlaAla 120
DB 954 CAAGACTGTGTTAAGCCCGCTCAGCTGCCCTCAGCGAGTGAAGTCTCTATCGAGCT 1013
QY 121 ProSerArgProGluProAsnArgCysGlySerCysArgLysArgValGlyLeuThrGly 140
DB 1014 CCTCTAGACCCGAGCCCAATGATGCGATCTTGACAGAGCGTGTGATTTGACAGCA 1073
QY 141 PheLysCysArgCysGlyAsnLeuTyrCysAlaLeuHisArgTyrSerAspLysHisThr 160
DB 1074 TTTAAGTGTGCTGTGGCAACCTCTACTGCGCTTTACATCGGACTCGAACAAACACT 1133
QY 161 CysThrTyrAspTyrLysAlaAlaGlyGlnGluAlaIleAlaLysAlaAsnProLeuVal 180
DB 1134 TGCACATATGACTACCAAGCCGACGGCAGAGCGATGCGAAGCTATCTCTTGTTC 1193
QY 181 ValAlaGluLysValValLysPhe 188
DB 1194 GTGGCCGAGAAAGTTGTCAAGTTT 1217

RESULT 4
US-10-716-089-4
; Sequence 4, Application US/10716089
; Publication No. US20040107463A1
; GENERAL INFORMATION:
; APPLICANT: COSTA E SILVA, OSWALDO DA
; APPLICANT: BOHNER, HANS J.
; APPLICANT: VAN THIELEN, NOCHA
; APPLICANT: CHEN, ROUYING
; TITLE OF INVENTION: TRANSCRIPTION FACTOR STRESS-RELATED PROTEINS AND
; TITLE OF INVENTION: METHODS OF USE IN PLANTS
; FILE REFERENCE: 16313-0030
; CURRENT APPLICATION NUMBER: US/10/716,089
; CURRENT FILING DATE: 2003-11-18
; PRIOR APPLICATION NUMBER: 60/196,001
; PRIOR FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: Patencin Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1531
; TYPE: DNA
; ORGANISM: Physcomitrella patens
US-10-716-089-4

Alignment Scores:
Pred. No.: 1,58e-101 Length: 1531
Score: 957.00 Matches: 187
Percent Similarity: 99.47% Conservative: 0

Best Local Similarity: 99.47% Mismatches: 1
Query Match: 97.26% Indels: 1
DB: 7 Gaps: 0
US-10-716-089-20 (1-188) x US-10-716-089-4 (1-1531)

QY 1 MetAlaThrGluArgValSerGlnGluThrThrSerGlnAlaProGlnGlu 20
DB 655 ATGGCCACCGAGCGTGTCTCTCAGAGAGACCTCCGAGGCCCTTAGGGCTCAGTTAG 714
QY 21 CysLysAsnLeuCysGlyPhePheGlySerGlnAlaThrMetGlyLeu 40
DB 715 TGCAGAACTTTGCGGCTTCTTCCGACGACGACTACCATGGGGTGTGCTGCAAGTGC 774
QY 41 TyrArgGluThrValMetGlnAlaLysMetThrAlaLeuAlaGluGlnAla 60
DB 775 TACCGAGACAGTCATGACGCG-AAGATGACGGCTTTAGCTGAGCAAC 833
QY 61 AlaGlnAlaThrSerAlaThrAlaAlaValGlnProProAlaProValHisGluThr 80
DB 834 GCTCAGGCGACATCTGCCACAGCTGCTGTTCAAGCCCCGCTCTGTACATGAGACC 893
QY 81 LysLeuThrCysGluValGluArgThrMetIleValProHisGlnSerSerSerTyrGln 100
DB 894 AAGCTCACATGCGAGGTGAGAGAACATGATGTGCGCCATCAATCTTCCAGCTATCA 953
QY 101 GlnAspLeuValThrProAlaAlaAlaProGlnAlaValLysSerSerIleAlaAla 120
DB 954 CAAGACTGTGTTAAGCCCGCTCAGCTGCCCTCAGCGAGTGAAGTCTCTATCGAGCT 1013
QY 121 ProSerArgProGluProAsnArgCysGlySerCysArgLysArgValGlyLeuThrGly 140
DB 1014 CCTCTAGACCCGAGCCCAATGATGCGATCTTGACAGAGCGTGTGATTTGACAGCA 1073
QY 141 PheLysCysArgCysGlyAsnLeuTyrCysAlaLeuHisArgTyrSerAspLysHisThr 160
DB 1074 TTTAAGTGTGCTGTGGCAACCTCTACTGCGCTTTACATCGGACTCGAACAAACACT 1133
QY 161 CysThrTyrAspTyrLysAlaAlaGlyGlnGluAlaIleAlaLysAlaAsnProLeuVal 180
DB 1134 TGCACATATGACTACCAAGCCGACGGCAGAGCGATGCGAAGCTATCTCTTGTTC 1193
QY 181 ValAlaGluLysValValLysPhe 188
DB 1194 GTGGCCGAGAAAGTTGTCAAGTTT 1217

RESULT 5
US-10-425-114-26146
; Sequence 26146, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack B.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecu
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Impro
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 26146
; LENGTH: 945
; TYPE: DNA
; ORGANISM: Zea mays
; OTHER INFORMATION: Clone ID: LIB4073-024-B8_FLI
US-10-425-114-26146

Alignment Scores:
Pred. No.: 1,35e-39 Length: 945

Score:	425.50	Matches:	83
Percent Similarity:	61.67%	Conservative:	28
Best Local Similarity:	46.11%	Mismatches:	56
Query Match:	43.24%	Indels:	13
DB:	7	Gaps:	2
US-10-716-089-20 (1-188) x US-10-425-114-26146 (1-945)			
QY	8	GIngluThrThrSerGlnAlaProGluGlyProValMetCysblyAsnLeuCyseGlyPhe	27
DB	112	AAGGAGACTGGATGCTGAGGACCTGAGGAGCCATCTTTGGATCAATTAACCTGGCTTC	171
QY	28	PheGlySerGlnAlaThrMetGlyLeuCyseSerLysCysTyraArgLuthrValMetGln	47
DB	172	TTCCGACGGCAGCTACCACTGAACATGCTCCAACTGCCACAGAGATGATTAAG	231
QY	48	AlaIysMetThrAlaLeuAlaGluGlnAlaThrGlnAlaAlaGlnAlaThrSerAlaThr	67
DB	232	CAGAGACAGGCCAAGCTGGCT-----GCCCTCTATGACAGCATGCTCAATGGCAAC	285
QY	68	AlaAlaAlaValGlnProProAlaProValHisGluThrLysLeuThrCysGluValGlu	87
DB	286	GATGCTGTCAATGAACA-----GTTGTTCT	312
QY	88	ArgThrMetIleValProHisGlnSerSerSerTyrgInGlnAspLeuValThrProAla	107
DB	313	GGCAACACAGTGTGCTGCTCTCCAACTCCAGTTGCAAAACAATGACGTGACGCCCT	372
QY	108	AlaAlaAlaProGlnAlaValLysSerSerIleAlaAlaProSerArgProGluProAla	127
DB	373	GATGTGCTGACCTTGCAGAGGGGGCGGTGATCTCCAAAGGAGTAGGCCGAAC	432
QY	128	ArgCysGlySerCysArgLysArgValGlyLeuThrGlyPheLysCysArgCysGlyAsn	147
DB	433	CGGTGACAGCACTTGCAAGAGGGTGTGACTTACAGATTCAACTGCCGTGTGGAAAC	492
QY	148	LeuTyrcysAlaLeuHisArgTySerAspLysHisThrCysThrTyraSPTyrLysAla	167
DB	493	TTGTACTGTGCATCTGCACCGCTACTCCAGACAGCACTGCAAGTTCTGACTGTGCACT	552
QY	168	AlaGlyGlnGluAlaIleAlaLysAlaAsnProLeuValAlaGluLysValValLys	187
DB	553	GCTGCTAGGATGATGCTCCAAAGCTTAATCCAGTGTGAAGCGGACAGACTGCACAAAG	612
RESULT 5			
US-10-425-115-71431			
; Sequence 71431, Application US/10425115			
; Publication No. US20040214272A1			
; GENERAL INFORMATION:			
; APPLICANT: Kovalic, David J.			
; APPLICANT: Zhou, Yihua			
; APPLICANT: Cao, Yongwei			
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with			
; FILE REFERENCE: 38-21(5322)B			
; CURRENT APPLICATION NUMBER: US/10/425,115			
; CURRENT FILING DATE: 2003-04-28			
; NUMBER OF SEQ ID NOS: 369326			
; SEQ ID NO 71431			
; LENGTH: 947			
; TYPE: DNA			
; ORGANISM: Zea mays			
; FEATURE:			
; OTHER INFORMATION: Clone ID: MRT4577_16514C.1			
US-10-425-115-71431			
Alignment Scores:			
Pred. No.:	1,356-39	Length:	947
Score:	425.50	Matches:	83
Percent Similarity:	61.67%	Conservative:	28
Best Local Similarity:	46.11%	Mismatches:	56
Query Match:	43.24%	Indels:	13

DB:	8	Gaps:	2
US-10-716-089-20 (1-188) x US-10-425-115-71431 (1-947)			
QY	8	GIngluThrThrSerGlnAlaProGluGlyProValMetCysblyAsnLeu	27
DB	114	AAGGAGACTGGATGCTGAGGACCTGAGGAGCCATCTTTGGATCAATTAACCTGGCTTC	173
QY	28	PheGlySerGlnAlaThrMetGlyLeuCyseSerLysCysTyraArgLuthrValMetGln	47
DB	174	TTCCGACGGCAGCTACCACTGAACATGCTCCAACTGCCACAGAGATGATTAAG	233
QY	48	AlaIysMetThrAlaLeuAlaGluGlnAlaThrGlnAlaAlaGlnAlaThrSerAlaThr	67
DB	234	CAGAGACAGGCCAAGCTGGCT-----GCCCTCTATGACAGCATGCTCAATGGCAAC	287
QY	68	AlaAlaAlaValGlnProProAlaProValHisGluThrLysLeuThrCysGluValGlu	87
DB	288	GATGCTGTCAATGAACA-----GTTGTTCT	314
QY	88	ArgThrMetIleValProHisGlnSerSerSerTyrgInGlnAspLeuValThrProAla	107
DB	315	GGCAACACAGTGTGCTGCTCTCCAACTCCAGTTGCAAAACAATGACGTGACGCCCT	374
QY	108	AlaAlaAlaProGlnAlaValLysSerSerIleAlaAlaProSerArgProGluProAla	127
DB	375	GATGTGCTGACCTTGCAGAGGGGGCGGTGATCTTCCAAAGGAGAGCT	434
QY	128	ArgCysGlySerCysArgLysArgValGlyLeuThrGlyPheLysCysArgCysGlyAsn	147
DB	435	CGGTGACAGCACTTGCAAGAGGGTGTGACTTACAGATTCAACTGCCGTGTGGAAAC	494
QY	148	LeuTyrcysAlaLeuHisArgTySerAspLysHisThrCysThrTyraSPTyrLysAla	167
DB	495	TTGTACTGTGCATCTGCACCGCTACTCCAGACAGCACTGCAAGTTCTGACTGTGCACT	554
QY	168	AlaGlyGlnGluAlaIleAlaLysAlaAsnProLeuValAlaGluLysValValLys	187
DB	555	GCTGCTAGGATGATGCTCCAAAGCTTAATCCAGTGTGAAGCGGACAGACTGCACAAAG	614
RESULT 7			
US-10-767-701-15574			
; Sequence 15574, Application US/10767701			
; Publication No. US20040172684A1			
; GENERAL INFORMATION:			
; APPLICANT: Kovalic, David J.			
; APPLICANT: Zhou, Yihua			
; APPLICANT: Cao, Yongwei			
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules			
; FILE REFERENCE: 38-21(5335)B			
; CURRENT APPLICATION NUMBER: US/10/767,701			
; CURRENT FILING DATE: 2004-01-29			
; NUMBER OF SEQ ID NOS: 63128			
; SEQ ID NO 15574			
; LENGTH: 1142			
; TYPE: DNA			
; ORGANISM: Sorghum bicolor			
; FEATURE:			
; OTHER INFORMATION: Clone ID: SORBI-28WAY03-CLUS087_1			
US-10-767-701-15574			
Alignment Scores:			
Pred. No.:	1,726-39	Length:	1142
Score:	425.50	Matches:	83
Percent Similarity:	61.67%	Conservative:	28
Best Local Similarity:	46.11%	Mismatches:	56
Query Match:	43.24%	Indels:	13
DB:	7	Gaps:	2
US-10-716-089-20 (1-188) x US-10-767-701-15574 (1-1142)			
QY	8	GIngluThrThrSerGlnAlaProGluGlyProValMetCysblyAsnLeu	27

```
Db 112 AAGAGACTGCATGCCAGGCACCTGAGGACCCATCTTGGATCAATTAATCGGCTTC 231
      ::::::::::|
Qy 28 PheGlySerGlnAlaThrMetGlyLeuCySerIysCyTyrArgGluThrValMetGln 47
      ::::::::::|
Db 232 TTGGCAGCGGAGCTACATGTAACATGTGCTCCAAAGTCCACAGAGATGATTAATGAAG 291
      ::::::::::|
Qy 48 AlaIysMetThrAlaLeuAlaGluGlnAlaThrGlnAlaAlaGlnAlaThrSerAlaThr 67
      ::::::::::|
Db 232 CAGGAGCAGGCCAAGCTGGCT-----GCCCTCTTATGCAGACAGCATCTGCAATGGCAAC 345
      ::::::::::|
Qy 68 AlaAlaAlaValAlaGlnProProAlaProValHisGluThrIysLeuThrCySerGluValGlu 87
      ::::::::::|
Db 346 GATCGCTGTCATGGAACCA-----GTTGTTGCT 372
      ::::::::::|
Qy 88 ArgThrMetIleValProHisGlnSerSerTyrGlnGlnAspLeuValThrProAla 107
      ::::::::::|
Db 373 GCGAACACACAGTGGTGGCTGCTGCTCCAAATCGATTGCACAAACATGAACGTGACGCCGCT 432
      ::::::::::|
Qy 108 AlaAlaAlaProGlnAlaValIysSerSerIleAlaAlaProSerArgProGluProAsn 127
      ::::::::::|
Db 433 GATGTTGCTGCAGCTACGAGGGGGGGGGGTGATCTCCMAAGGAGATGAGCGCGAAC 492
      ::::::::::|
Qy 128 ArgCySerGlySerCyArgIysArgValGlyLeuThrGlyPheIysCyBArgCySerGluAsn 147
      ::::::::::|
Db 493 CCGTCCAGACACTTGCAGAGAGAGGGTTGGACTTACAGATTCAACTGCCGGTGTGGAGAC 552
      ::::::::::|
Qy 148 LeuTyrCyValaLeuHisArgTyrSerAspIysHisThrCySerThrTyrAspTyrIysAla 167
      ::::::::::|
Db 553 TTGTACTGTGCACCTGCACCGCTACTCCGACAGACAGACTGCAGATTGCACTATGCGACT 612
      ::::::::::|
Qy 168 AlaGlyGlnGluAlaIleAlaIysAlaAsnProLeuValValAlaGluIysValIys 187
      ::::::::::|
Db 613 GCTGCTAGGAGATGCATTGCGCAAGGCTAATCCAGTGTGAAGCGGACAGACTCGACAAAG 672
      ::::::::::|
```

RESULT 8

```
US-10-437-963-21113
; Sequence 21113, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovacic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; NUMBER OF SEQ ID NOS: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 21113
; LENGTH: 2671
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_26415C.1
US-10-437-963-21113
```

```
Alignment Scores:
Pred. No.: 2,586-38 Length: 2671
Score: 419.50 Matches: 82
Percent Similarity: 60.09% Conservative: 26
Best Local Similarity: 45.56% Mismatches: 59
Query Match: 42.63% Indels: 13
DB: 7 Gaps: 2
```

US-10-716-089-20 (1-188) x US-10-437-963-21113 (1-2671)

Qy 8 GlnGluThrThrSerGlnAlaProGluGlyProValMetCyBlyAsnLeuCySerGlyPhe 27

```
Db 1457 AAGAGACTGCATGCCAGGCACCTGAGGACCCATCTTGGATCAATTA 1516
      ::::::::::|
Qy 28 PheGlySerGlnAlaThrMetGlyLeuCySerIysCyTyrArgGluThr 47
      ::::::::::|
Db 1517 TTGGCAGCGGAGCTACATGTAACATGTGCTCCAAAGTCCACAGAGAT 1576
      ::::::::::|
Qy 48 AlaIysMetThrAlaLeuAlaGluGlnAlaThrGlnAlaAlaGlnAlaThr 67
      ::::::::::|
Db 1577 CAGGAGCAGGCCAAGCTGGCA-----GCCCTCTTATGCAGACAGATTG 1630
      ::::::::::|
Qy 68 AlaAlaAlaValAlaGlnProProAlaProValHisGluThrIysLeuThr 87
      ::::::::::|
Db 1631 GATTCGGGAGAGAACCAATTATCTGCTGCTACGCTGAAGTACGTTTC 1690
      ::::::::::|
Qy 88 ArgThrMetIleValProHisGlnSerSerTyrGlnGlnAspLeu 107
      ::::::::::|
Db 1691 GTGAAGACGCTGTT-----G 1717
      ::::::::::|
Qy 108 AlaAlaAlaProGlnAlaValIysSerSerIleAlaAlaProSerArgP 127
      ::::::::::|
Db 1718 GAGATTGCTGTCCTAGTACGAGGGGGGTACAGGTGAACCCCAAGGGAGG 1777
      ::::::::::|
Qy 128 ArgCySerGlySerCyArgIysArgValGlyLeuThrGlyPheIysCyB 147
      ::::::::::|
Db 1778 CCGTCCAGACACTTGCAGAGAGAGGGTTGGACTTACAGATTCAACTGCC 1837
      ::::::::::|
Qy 148 LeuTyrCyValaLeuHisArgTyrSerAspIysHisThrCySerThrTyr 167
      ::::::::::|
Db 1838 TTGTACTGTGCACCTGCACCGCTACTCCGACAGACAGACTGCAGATTTC 1897
      ::::::::::|
Qy 168 AlaGlyGlnGluAlaIleAlaIysAlaAsnProLeuValValAlaGlu 187
      ::::::::::|
Db 1898 GCTGCTAGGAGATGCATTGCGCAAGGCTAATCCAGTGTGAAGCGGAGAG 1957
      ::::::::::|
```

RESULT 9

```
US-10-425-114-27574
; Sequence 27574, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovacic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules
; FILE REFERENCE: 38-21(53131)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 27574
; LENGTH: 1032
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4732-053-A11_FLI
US-10-425-114-27574
```

```
Alignment Scores:
Pred. No.: 1,716-38 Length: 1032
Score: 416.50 Matches: 85
Percent Similarity: 62.09% Conservative: 28
Best Local Similarity: 46.70% Mismatches: 56
Query Match: 42.33% Indels: 13
DB: 7 Gaps: 4
```

US-10-716-089-20 (1-188) x US-10-425-114-27574 (1-1032)

Qy 8 GlnGluThrThrSerGlnAlaProGluGlyProValMetCyBlyAsnLeu 27

Db 134 AAGAGCTGGGTGCAGAGCCCGAGGAGCCCATCTCTGATCAATTA 193

```
Qy      28 PheGlySerGlnAlaThrMetGlyLeuCySerIleCysTyrrArgGluThrValMetCln 47
Db      194 TTTCGACGGCGGACCAATGATGCTTAAGTCCACAGGAGATGATTAACGAG 253
Qy      48 AlaIysMetThrAlaLeuAlaGluGlnAlaThrGlnAlaAlaGlnAlaThrSerAlaThr 67
Db      254 CAGATGACGCGGACAGCTGGCT-----GCTTCCTGATTTGACAGCATGTTGAACGGCAGC 307
Qy      68 AlaAlaAlaValGlnProProAlaProValHisGluThrLysLeuThrCysGluValGlu 87
Db      308 GACGCCGCTGACGAGCCGCTTGTGCTGCTGACGACACAG-----GTACTA 352
Qy      88 ArgThrMetIleValProHisGlnSerSerSerTyrrGlnGlnAspLeuValThrProAla 107
Db      353 GCTGTGCTCCCAAGTCAGATTGCAACATGAACTGCGACAG-----CCGCT 400
Qy      108 -----AlaAlaAlaProGlnAlaValLysSerSerIleAlaAlaProSerArgProGlu 125
Db      401 GATGTGCGGACCCAGGAGGGGGGTGGCGGCGGATCTCCAAAGGGGGAGGTAGG 460
Qy      126 ProIleArgCysGlySerCysArgLysArgValGlyLeuThrGlyPheLysCysArgCys 145
Db      461 CCGAACCGGTCCAGCGCTGCGAGAGAGGGTGGCTCACGGAGATTCACTGCCGGTGT 520
Qy      146 GlyAsnLeuTyrrCysAlaLeuHisArgTyrrSerAspLysHisThrCysThrTyrAspTyr 165
Db      521 GGGAACTTGACTGATGGCGCTCCACCGCTACTCTCGCAAGACGAGCTGAAAGTTGCACCTAC 580
Qy      166 LysAlaAlaGlyGlnGlnAlaIleAlaLysAlaAsnProLeuValValaGluLysVal 185
Db      581 CGGACTGCTGCACGGGAGCGCATTTCCAAAGGCTATCCGGGTGTAAAGCAGACAGCTTC 640
Qy      186 ValLys 187
Db      641 GACCAAG 646

RESULT 10
US-10-437-963-37809/c
/ Sequence 37809, Application US/10437963
/ Publication No. US20040123343A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ APPLICANT: Wu, Wei
/ APPLICANT: Boukharov, Andrey A.
/ APPLICANT: Barbazuk, Brad
/ APPLICANT: Li, Ping
/ TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
/ FILE REFERENCE: 38-21(53221)B
/ CURRENT APPLICATION NUMBER: US/10/437,963
/ CURRENT FILING DATE: 2003-05-14
/ NUMBER OF SEQ ID NOS: 204966
/ SEQ ID NO 37809
/ LENGTH: 1171
/ TYPE: DNA
/ ORGANISM: Oryza sativa
/ FEATURE:
/ OTHER INFORMATION: Clone ID: PAT_MRT4530_41503C.1
US-10-437-963-37809

Alignment Scores:
Pred. No.: 2,01e-38 Length: 1171
Score: 416.50 Matches: 82
Percent Similarity: 57.29% Conservative: 28
Best Local Similarity: 42.71% Mismatches: 47
Query Match: 42.33% Indels: 35
DB: 7 Gaps: 3
US-10-716-089-20 (1-188) x US-10-437-963-37809 (1-1171)
```

```
Qy      8 GlnGluThrThrSerGlnAlaProGluGlyProValMetCysLysAsnL 27
Db      890 AAGGAGGCTGCTGCCACAGCCCGAGAGGCCAATCTTTCATCAATAT 831
Qy      28 PheGlySerGlnAlaThrMetGlyLeuCySerIleCysTyrrArgGluThr 47
Db      830 TTTCGACGGCGGACCAATGATGCTTAAGTCCACAGGAGATGATTAACGAG 771
Qy      48 AlaIysMetThrAlaLeuAlaGluGlnAlaThrGlnAla----- 60
Db      770 GAGGAGCAGGCCAAGCTTGTGCTGCTGCTCCATGATGATGATTCATG 711
Qy      61 -----AlaGlnAlaThrSerAlaThrAlaAlaValG 75
Db      710 GGGAGAGACATATTGTTGCTGCCAGTGTAGCAGCGCGGTGCGGTG- 663
Qy      76 ProValHisGluThrLysLeuThrCysGluValGluArgThrMetIleV 95
Db      662 -----GCTCAGGTTCGAGCGCAAGACGCTCG 636
Qy      96 SerSerSerTyrrGlnGlnAspLeuValThrProAlaAlaAlaProG 115
Db      635 -----GTGCAGCTCTACCGATGTCGGGCA 600
Qy      116 SerSerIleAlaAlaProSerArgProGluProAsnArgCysGlySerC 135
Db      599 GTTCTGTAGTCCCAAGCTGCAAGAGGGCCGAACCGGTGCGTCACTC 540
Qy      136 ValGlyLeuThrGlyPheLysCysArgCysGlyAsnLeuTyrrCysAlaL 155
Db      539 GTTGAGGCTGACGGGATTCACCTGCCGGGCGGTAAACATGATCTGGCGCT 480
Qy      156 SerAspLysHisThrCysThrTyrAspTyrLysAlaAlaGlyGlnAla 175
Db      479 TCCGACCAAGCATGAATGCCAGTTCGACTACCGAGCTGCGGCTAGGGATG 420
Qy      176 AlaAsnProLeuValValaGluLysValLys 187
Db      419 GCCAACCAAGTGTAAAGCTGAGAACTCCAGCAAG 384

RESULT 11
US-10-437-963-37808/c
/ Sequence 37808, Application US/10437963
/ Publication No. US20040123343A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ APPLICANT: Wu, Wei
/ APPLICANT: Boukharov, Andrey A.
/ APPLICANT: Barbazuk, Brad
/ APPLICANT: Li, Ping
/ TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Mo
/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Impro
/ FILE REFERENCE: 38-21(53221)B
/ CURRENT APPLICATION NUMBER: US/10/437,963
/ CURRENT FILING DATE: 2003-05-14
/ NUMBER OF SEQ ID NOS: 204966
/ SEQ ID NO 37808
/ LENGTH: 1249
/ TYPE: DNA
/ ORGANISM: Oryza sativa
/ FEATURE:
/ OTHER INFORMATION: Clone ID: PAT_MRT4530_41502C.1
US-10-437-963-37808

Alignment Scores:
Pred. No.: 2.18e-38 Length: 1249
Score: 416.50 Matches: 82
Percent Similarity: 57.29% Conservative: 28
Best Local Similarity: 42.71% Mismatches: 47
```


Pred. No.: 1.77e-38 Length: 776
Score: 415.00 Matches: 85
Percent Similarity: 61.88% Conservative: 27
Best Local Similarity: 46.96% Mismatches: 57
Query Match: 42.17% Indels: 12
DB: 7 Gaps: 4

US-10-716-089-20 (1-188) x US-10-425-114-28524 (1-776)

QY 8 GlnGluThrThrSerGlnAlaProGluGlyProValMetCysValAsnLeuCysGlyPhe 27
DB 133 AAGAGGCTGGGTGTCAGAGCCCCCGAGGAGACCCATCTCTGATCAATAACTGTCCTTC 192
QY 28 PheGlySerGlnAlaThrMetGlyLeuCysSerIlyCysGlyTyrArgGluThrValMetGln 47
DB 193 TTCGACAGCGCGGACCATGACATGTGCTTTAGTGCCACAGAGATGATTAACGAAG 252
QY 48 AlalyMetThrAlaLeuAlaGluGlnAlaThrGlnAlaAlaGlnAlaThrSerAlaThr 67
DB 253 CAGGATCAGGCCAAGCTGGCT-----GCTCTCTATCGACAGCATGTCGTAACGGCAGC 306
QY 68 AlAlaAlaValAlaGlnProProAlaProValHisGluThrIlySleuthrCysGluValGlu 87
DB 307 GACGCCGTATGAGCCGCTGTTGCTGCGACCAACAG-----GTAGTA 351
QY 88 ArgThrMetIleValProHisGlnSerSerIyGlnGlnApyLeuValThrProAla 107
DB 332 GCTGTGCCCAAGTGCAGTTGCAACATGAACGTGCAGCA-----CCCTCT 399
QY 108 ---AlAlaAlaProGlnAlaValIlySerSerIleAlaAlaProSerArgProGluPro 126
DB 400 GATGTGGCCGAGCCAGGAGGGGTGGCGCGATCTCCAAAGGGGAGTAGGGCGC 459
QY 127 AsnArgCysGlySerCysArgIlyArgValGlyLeuThrIlyPheIlyCysArgCysGly 146
DB 460 AACCGGTGACGCGCTGCAGAGAGGGTGGCTCACGGGATTCACGCGGTGTGG 519
QY 147 AsnLeuTyrCysAlaLeuHisArgTyrSerAspIlyValHisThrCysThrIlyAspTyrIly 166
DB 520 AACTGTGACTGGCGCTCCACCGCTACTCCGACAGCAAGCTGCAAGTTCACACGG 579
QY 167 AlAlaGlyGlnGlnAlaIleAlaIlyAlaAsnProLeuValAlaGluIlyValAla 186
DB 580 ACTGCTGCAGGAGCGCATTCGCAAGCTATCCGGTGTGAAGCAGACAACTGCAC 639
QY 187 Lys 187
DB 640 AAG 642

RESULT 14

US-10-425-114-22082
; Sequence 22082, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 22082
; LENGTH: 821
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3356-040-B2_FLI
US-10-425-114-22082

Alignment Scores:
Pred. No.: 1.9e-38 Length: 821
Score: 415.00 Matches: 85
Percent Similarity: 61.88% Conservative: 27
Best Local Similarity: 46.96% Mismatches: 57
Query Match: 42.17% Indels: 12
DB: 7 Gaps: 4

US-10-716-089-20 (1-188) x US-10-425-114-22082 (1-821)

QY 8 GlnGluThrThrSerGlnAlaProGluGlyProValMetCysValAsnLeu 27
DB 113 AAGAGGCTGGGTGTCAGAGCCCCCGAGGAGACCCATCTCTGATCAATAACTGTCCTTC 172
QY 28 PheGlySerGlnAlaThrMetGlyLeuCysSerIlyCysGlyTyrArgGluThrValMetGln 47
DB 173 TTCGACAGCGCGGACCATGACATGTGCTTTAGTGCCACAGAGATGATTAACGAAG 232
QY 48 AlalyMetThrAlaLeuAlaGluGlnAlaThrGlnAlaAlaGlnAlaThrSerAlaThr 67
DB 233 CAGGATCAGGCCAAGCTGGCT-----GCTCTCTATCGACAGCATGTCGTAACGGCAGC 286
QY 68 AlAlaAlaValAlaGlnProProAlaProValHisGluThrIlySleuthrCysGluValGlu 87
DB 287 GACGCCGTATGAGCCGCTGTTGCTGCGACCAACAG-----GTAGTA 351
QY 88 ArgThrMetIleValProHisGlnSerSerIyGlnGlnApyLeuValThrProAla 107
DB 332 GCTGTGCCCAAGTGCAGTTGCAACATGAACGTGCAGCA-----CCCTCT 399
QY 108 ---AlAlaAlaProGlnAlaValIlySerSerIleAlaAlaProSerArgProGluPro 126
DB 380 GATGTGGCCGAGCCAGGAGGGGTGGCGCGATCTCCAAAGGGGAGTAGGGCGC 439
QY 127 AsnArgCysGlySerCysArgIlyArgValGlyLeuThrIlyPheIlyCysArgCysGly 146
DB 440 AACCGGTGACGCGCTGCAGAGAGGGTGGCTCACGGGATTCACGCGGTGTGG 499
QY 147 AsnLeuTyrCysAlaLeuHisArgTyrSerAspIlyValHisThrCysThrIlyAspTyrIly 166
DB 500 AACTGTGACTGGCGCTCCACCGCTACTCCGACAGCAAGCTGCAAGTTCACACGG 559
QY 167 AlAlaGlyGlnGlnAlaIleAlaIlyAlaAsnProLeuValAlaGluIlyValAla 186
DB 560 ACTGCTGCAGGAGCGCATTCGCAAGCTATCCGGTGTGAAGCAGACAACTGCAC 619
QY 187 Lys 187
DB 620 AAG 622

RESULT 15

US-10-425-114-21492
; Sequence 21492, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 21492
; LENGTH: 826
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3356-040-B2_FLI
US-10-425-114-21492

! OTHER INFORMATION: Clone ID: L183354-001-F10_FLI
US-10-425-114-21492

Alignment Scores:

Pred. No.:	1.92e-38	Length:	826
Score:	415.00	Matches:	85
Percent Similarity:	61.88%	Conservative:	27
Best Local Similarity:	46.96%	Mismatches:	57
Query Match:	42.17%	Indels:	12
DB:	7	Gaps:	4

US-10-716-089-20 (1-188) x US-10-425-114-21492 (1-826)

```

QY      8 GInGUThThrSerGlnAlaProGlnUGlYPRoValMeCYeLyBaenLeuCYsgLYpHe 27
      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB      8 AAGGAGCTGGGTGGCCAGGCCCCGAGGAGCCCATCTCTGCATCAATACTGTGGCTTC 67
      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY     28 PheGlySerGlnAlaThrMetGlyLeuCYeSerLYeCYeTYzArgGluThrValMetGln 47
      |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB     68 TTCGGCAGCCCGCGGACATGAACTGTGCTTAAGTGCACAGAGATGATACGAAAG 127
      |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY     48 AlaLyMeetHrAlaLeuAlaGlnAlaThrGlnAlaAlaGlnAlaThrSerAlaThr 67
      |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB    128 CAGGATCAGGCCAAGCTGGCT-----GCCTCTATGCACAGCATCGTAACGGCAGC 181
      |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY     68 AlaAlaAlaValGlnProProAlaProValHisGluThrLYeLeuThrCYsgLYuValGlu 87
      |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB    182 GACGCCGTCATGGAGCGGTGTGGTGTGGCAGCACACG-----GTAGTA 226
      |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY     88 ArgThrMetIleValProHisGlnSerSerTYrGlnGlnAspLeuValThrProAla 107
      |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB    227 GCTGTGCCCAAGTCCAGTTGCCAAACATGACGTGCAGCAG-----CCGGCT 274
      |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY    108 ---AlaAlaAlaProGlnAlaValLYeSerSerIleAlaAlaProSerArgProGlnPro 126
      |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB    275 GATGTTGCCGAGCCAGCGAGGGGGGTGGCGCATCTCCAAAGGGGGAAGGTAGGGCGG 334
      |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY    127 AsnArgCYeGlySerCYeArgLYeArgValGlyLeuThrGlyPheLYeCYeArgCYeGly 146
      |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB    335 AACCGGTGACGCCCTGCAGGAGAGAGGGGTGGCTCACGGGATTCAACTGCCGTGTGGG 394
      |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY    147 AsnLeuTYrCYeAlaLeuHisArgTYrSerAspLYeHisThrCYeThrTYrAspTYrLYe 166
      |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB    395 AACTTGTAAGTCCCGCTCCACCGCTACTCCGACAGACAGACTGCAGATTCACTACCGG 454
      |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY    167 AlaAlaGlyGlnGlnAlaIleAlaIleAlaLYeAlaAsnProLeuValAlaGluLYeVal 186
      |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB    455 ACTGCTGCCAGGAGCGCATTTGCCAAGGCTAATCCGTGTGTGAAGGCAGCAAGCTCGAC 514
      |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY    187 LYe 187
      |||||
DB    515 AAG 517
      |||||

```

Search completed: December 8, 2005, 18:57:54
Job time : 795 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 8, 2005, 16:44:44 ; Search time 249 Seconds

(without alignments)
282.268 Million cell updates/sec

Title: US-10-716-089-20

Perfect score: 984

Sequence: 1 MATEVSEGTSSQAEQGVW.....GGEALAKANPLVAKVKYKF 188

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 3392430 seqs, 186927314 residues

Total number of hits satisfying chosen parameters: 6784860

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+g2n.model -DEV=slp
-Q=/cgn2_1/USPTO_spool_p/US10716089/runat_05122005_094345_13229/app_query.fasta_1.327
-DB=Published Applications NA New -OPMT=fastap -SUFFIX=g2n.rnpbn -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blowsum62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pro -NOR=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US10716089_@cgn2_1.184_@runat_05122005_094345_13229
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA New.*

1: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
2: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
5: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq2.*
9: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq3.*
10: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	189.5	19.3	1067	US-10-750-185-47138	Sequence 47138, A
C 2	180	18.3	1563	US-10-750-185-52649	Sequence 52649, A
C 3	90.5	9.2	864	US-10-467-657-1869	Sequence 1869, Ap
C 4	88.5	9.0	2133	US-11-045-802-10	Sequence 10, Appl
5	86.5	8.8	3950	US-10-821-234-253	Sequence 253, App
6	86	8.7	2089	US-10-750-185-60675	Sequence 60675, A
7	86	8.7	3513	US-10-467-657-7841	Sequence 7841, Ap
C 8	86	8.7	7893	US-11-186-731-3	Sequence 3, Appl

C 9	86	8.7	8106	7	US-11-186-731-1	Seq
C 10	86	8.7	23907	7	US-11-186-731-6	Seq
C 11	86	8.7	24120	7	US-11-186-731-4	Seq
C 12	85	8.6	13323	7	US-11-087-100-21	Seq
C 13	85	8.6	6177	7	US-11-087-100-3	Seq
C 14	83	8.4	2815	6	US-10-821-234-165	Seq
C 15	82.5	8.4	801	6	US-10-508-263-103	Seq
C 16	82.5	8.4	1076	6	US-11-102-240-45	Seq
C 17	82.5	8.4	2159	6	US-10-131-826A-87	Seq
C 18	82	8.3	1161	6	US-10-858-730-170	Seq
C 19	81	8.2	16816	7	US-11-121-086-3	Seq
C 20	80.5	8.2	1991	6	US-10-750-185-27360	Seq
C 21	80.5	8.2	2825	6	US-10-750-185-56778	Seq
C 22	80.5	8.2	3513	6	US-10-467-657-7841	Seq
C 23	80	8.1	1022	6	US-10-750-185-63872	Seq
C 24	80	8.1	8424	6	US-10-821-234-47	Seq
C 25	80	8.1	10373	6	US-10-821-234-64	Seq
C 26	79.5	8.1	2133	7	US-11-045-802-11	Seq
C 27	79.5	8.1	2302	7	US-11-110-082-20	Seq
C 28	79	8.0	1102	6	US-10-131-826A-299	Seq
C 29	79	8.0	1392	7	US-11-182-946-1	Seq
C 30	79	8.0	1511	6	US-10-750-185-37762	Seq
C 31	79	8.0	6058	6	US-10-770-726-17	Seq
C 32	79	8.0	168516	7	US-11-121-086-3	Seq
C 33	79	8.0	169495	7	US-11-121-086-61	Seq
C 34	78.5	8.0	1017	6	US-10-909-125-823	Seq
C 35	78.5	8.0	3334	6	US-10-821-234-152	Seq
C 36	78.5	8.0	8424	6	US-10-821-234-47	Seq
C 37	78.5	8.0	14342	6	US-10-821-234-303	Seq
C 38	78.5	8.0	78869	7	US-11-075-185-1	Seq
C 39	78.5	8.0	197096	7	US-11-121-086-107	Seq
C 40	78	7.9	407	7	US-11-108-112-174	Seq
C 41	78	7.9	1140	6	US-10-793-626-2809	Seq
C 42	78	7.9	1467	6	US-10-821-234-802	Seq
C 43	78	7.9	1675	6	US-10-750-185-41941	Seq
C 44	78	7.9	2260	7	US-11-186-284-120	Seq
C 45	78	7.9	2988	6	US-10-793-626-4132	Seq

ALIGNMENTS

RESULT 1
US-10-750-185-47138/C
; Sequence 47138, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 47138
; LENGTH: 1067
; TYPE: DNA
; ORGANISM: Bovine 19866881619274
US-10-750-185-47138
Alignment Scores:
Pred. No.: 7.86e-09
Score: 189.50
Percent Similarity: 52.24%
Best Local Similarity: 32.84%
Query Match: 19.26%
Length: 1067
Matches: 66
Conservative: 39
Mismatches: 61
Indels: 38

Oy	49	lywemethzrlaleumlaagluglnlaethglnalalaglnlaphserlathzrla	68
Db	18137	ctgattgcacagccttgggtgtgccctgcacccggcagcagccagcttgatctgtccacccac	18078
Oy	69	AlaAlaValGlnPro-ProAlaProValHisIuThnIyLeuthrCyGluValGluAr	88
Db	18077	tcggcttgcagatgcgcgcagactctcttccaaagtccggggggccgcacacagcagggcc	18018
Oy	88	grhmetilevalProHisInserSerSerTyrgInIasp-----LeuValth	105
Db	18017	agacgcgtgcgatgccacagatctcttcacccacagactctgtatnagctgtgcgt	17958
Oy	105	rProAlaAlaAlaProGlnAlaValIySerSerIleAlaAlaProSerArg----	123
Db	17957	gcttcgacagcagatcttcgcacacccagatctctccgcttcgcacacatcgaaaggcgcg	17898
Oy	124	-----ProGluProAsnArgCyGlySerCyArgIyAsnValGlyLeuthrGlyph	141
Db	17897	tcatccccctccacactgctggttcaggtcgatctctcagcttcattcatgtgtccgaac	17838
Oy	141	elysCys-----ArgCysGlyAsnLeu	148
Db	17837	acgtacgtacacggtatcgtggtccggagatcttc	17807

```

RESULT 11
US-11-186-731-4/c
; Sequence 4, Application US/1186731
; Publication No. US20050255521A1
; GENERAL INFORMATION:
; APPLICANT: Kapellier-Libermann, Rosana
; APPLICANT: Acton, Susan L.
; TITLE OF INVENTION: 59079 and 12599, Protein Kinase Family
; TITLE OF INVENTION: Members and Uses Therefor
; FILE REFERENCE: MP12001-047P1RCP1(M)
; CURRENT APPLICATION NUMBER: US/11/186,731
; CURRENT FILING DATE: 2005-07-21
; PRIOR APPLICATION NUMBER: US/10/077,130
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: 60/269201
; PRIOR FILING DATE: 2001-02-15
; NUMBER OF SEQ. ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 24120
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (1)...(71)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (72)...(23978)
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: (23979)...(24120)
US-11-186-731-4

Alignment Scores:
Pred. No.:      3.59e+03
Score:          86.00
Percent Similarity: 37.40%
Best Local Similarity: 28.24%
Query Match:    8.74%
              7
Gaps:          3
              3
Length:        24120
Matches:       37
Conservative:  12
Mismatch:      71
Indels:        11
Gaps:          3

```

US-10-716-089--20 (1-188) x US-11-186-731-4 (1-24120)

Oy 29 GlysSerGlnAlaIleMetCysGlyLeuCysSerLysCysPyrArgGluThrValMetGlnAla 48
 ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 18268 GGCGTTCATAGAGATGTGGTGAGGCTCCACCTGCACCTGTTCCCATCTTTCACAG 18209
 ||||| ||||| ||||| ||||| ||||| ||||| |||||

Oy 49 LysMetThrAlaLeuAlaGluGlnAlaThrGlnAlaAlaGlnAlaThrSerAlaTrpAla 68
 ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

Db      18208  CTGATGACGAGCTTGAGGTGTGTCGCCCGACGCGGACGAGCCAGCTTGAATG
QY      69   ALAAlaValGlnPro-ProAlaProValHnsgLutHnlyseUthrCys
Db      18148  TCGGCTGTGCACATTCGCGCCAGCTCTCTTTAAAGTCGGGGGCGCGCAC
QY      88   5ThrmEtIleValProHnEglnSerSerSertyrGlnGlnAsp-----
Db      18088  AGAGCTGTGTGATGCGACAGATCTCTTCAACCGACGAGCTTGTATTA
QY      105  rProAlaAlaAlaAlaProGlnAlaValIlysserSerIleAlaAlaPro
Db      18028  GCCTGCAGAGGTACTTGGCACCAGTCTCCCGCTCCGCTCCGCCACACT
QY      124  -----ProGlnProsmnArgCysEgIlySerCysArgIlyArgValGly
Db      17968  TCATCCCCCTCCACCGTCGTCGTTCAAGTGGATCTGTGCTACGTTCAATCA
QY      141  eLysCys-----ArgCysGlyAsnIleu 148
Db      17908  ACGTAGCTGACGGTATTCGGTGTGGGGAGATCTC 17878

```

```

RESULT 12
US-11-087-100-21
/ Sequence 21, Application US/11087100
/ Publication No. US2005026640A1
/ GENERAL INFORMATION:
/ APPLICANT: Metz, James
/ APPLICANT: Barclay, William
/ APPLICANT: Platt, James
/ APPLICANT: Kuner, Jerry
/ TITLE OF INVENTION: Nucleic Acid Molecule Encoding ORF of a
/ TITLE OF INVENTION: System and Uses Thereof
/ FILE REFERENCE: 2997-23
/ CURRENT APPLICATION NUMBER: US/11/087,100
/ CURRENT FILING DATE: 2005-03-21
/ PRIOR APPLICATION NUMBER: 09/231,899
/ PRIOR FILING DATE: 1999-01-14
/ PRIOR APPLICATION NUMBER: 60/284,066
/ PRIOR FILING DATE: 2001-04-16
/ PRIOR APPLICATION NUMBER: 60/298,796
/ PRIOR FILING DATE: 2001-06-15
/ PRIOR APPLICATION NUMBER: 60/323,269
/ PRIOR FILING DATE: 2001-09-18
/ NUMBER OF SEQ ID NOS: 37
/ SOFTWARE: PatentIn version 3.3
/ SEQ ID NO 21
/ LENGTH: 1323
/ TYPE: DNA
/ ORGANISM: Schizochytrium sp.
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (1)..(1323)
US-11-087-100-21

```

Alignment Scores:	
Pred. No.:	123
Score:	85.00
Percent Similarity:	39.29%
Best Local Similarity:	25.51%
Query Match:	8.64%
DB:	7
	Gaps: 10
	Matches: 27
	Conservative: 50
	Mismatches: 69
	Indels: 5
	Gaps: 10

US-10-716-089-20 (1-188) x US-11-087-100-21 (1-1323)

[illegible]

```

QY      47 GlnAlaLysMetThrAlaLeuAlaGluGlnAlaThrGlnAlaGlnAlaThrSerAla 66
      335 GAATGAAAAAGGTGGCAATGTCGGCTTGTGGCCCTCGACCGACCTGCTCT 394
QY      67 TTT-----AlaAlaAlaGlnProAlaPro 76
      395 ACCGTCAACGGTGTCCGCTCGCTCAAGAGGCGCTCGCCCTGAAGCTCA----- 448
QY      77 ValHisGluThrLysLeuThr---CysGluValGluArgThrMetIleValProHisGln 95
      449 -----AGAAAGCTCAATGACATGATGACATTAACAGCATGCGGACATCCACATCCT 502
QY      96 SerSerSerTyrGlnGlnAspLeuValThr----- 105
      503 ACA-----CCTGTCATTTGGCAACCTGTCGCCACGCGGCTGCTCGCACTGGG 553
QY      106 -----ProAlaAlaAlaAlaProGlnAlaValLysSerIleAlaAlaPro 121
      554 GCTTCAAGGAGCCCTCTTACGATCAGACGAGGCAACAACCTCCGCTACCGCTGCGCG 613
QY      122 Ser-Arg-----ProGluProAsnArgCysGlySerCysArgLysArgValGlyLeuTh 139
      614 ACCTCGGCAAGTACCTCTCGAGACCGGAGGTGATGCGCTGCGGAGTGTG 673
QY      139 rGlyPheLysCysArgCys-GlyAsnLeuTyrCysAlaLeuHisArgTyr----- 155
      674 ATCTCTGCGG---CACTGCGGAACCTTTACGTCAAGTCTCGCCCTTCAAGTGTCCA 730
QY      156 --SerAspLysHisThrCysThrTyrAspTyrLysAlaAlaGly 169
      731 CTTCCGATACCCCGCGCGCCAGCTTTGAAGCGCGCCGATGCG 774
      DB
      731 CTTCCGATACCCCGCGCGCCAGCTTTGAAGCGCGCCGATGCG 774

RESULT 13
US-11-087-100-3
; Sequence 3, Application US/11087100
; Publication No. US20050266440A1
; GENERAL INFORMATION:
; APPLICANT: Metz, James
; APPLICANT: Barclay, William
; APPLICANT: Platt, James
; APPLICANT: Kuner, Jerry
; TITLE OF INVENTION: Nucleic Acid Molecule Encoding ORPA of a PUPA Polyketide Synthase
; TITLE OF INVENTION: System and Uses Thereof
; FILE REFERENCE: 2997-23
; CURRENT APPLICATION NUMBER: US/11/087,100
; CURRENT FILING DATE: 2005-03-21
; PRIOR APPLICATION NUMBER: 09/231,899
; PRIOR FILING DATE: 1999-01-14
; PRIOR APPLICATION NUMBER: 60/284,066
; PRIOR FILING DATE: 2001-04-16
; PRIOR APPLICATION NUMBER: 60/298,796
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/323,269
; PRIOR FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 3
; LENGTH: 6177
; TYPE: DNA
; ORGANISM: Schizochytrium sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(6177)
; US-11-087-100-3

```

Alignment Scores:

```

Pred. No.:      830      Length:      6177
Score:          85.00     Matches:      50
Percent Similarity: 39.29%  Conservative: 27
Best Local Similarity: 25.51%  Mismatches: 69
Query Match:      8.64%     Indels:      50
DB:              7         Gaps:      10

```

```

US-10-716-089-20 (1-188) x US-11-087-100-3 (1-6177)
QY      10 ThrHisSerGlnAlaProGluLysPro-----ValMetCysLys/
      1610 TCGACTTCAGGCGCTCCGACGAGCCCATGACCCCTGAAGACATGCTCC
QY      27 PhePheGlySerGlnAlaThrMetGlyLeuCysSerLysCysTyrArg
      1670 TTCTGCGCTGACCAACCATTTGACCGGCGCATCTCG-----
QY      47 GlnAlaLysMetThrAlaLeuAlaGluGlnAlaThrGlnAlaAlaGln
      1712 GAATGAAAAAGGTGGCAATGTCGGCTTGTGGCCCTCGACCG
      DB
      67 Thr-----AlaAlaAlaValGln
      1772 ACCGTCAACGGTGTCCGCTCGCTCAAGAGGCGGTGCGCCCTGAAGC
QY      77 ValHisGluThrLysLeuThr---CysGluValGluArgThrMetIleVal
      1826 -----AGAAAGCTCAATGACATGATGACATTAACGACATGCGGAC
QY      96 SerSerSerTyrGlnGlnAspLeuValThr-----
      1880 ACA-----CCTGTCATTTGGCAACCTGTCGCCACGCGCTTC
      DB
      106 -----ProAlaAlaAlaAlaProGlnAlaValLysSerIle
      1931 GCTTCAAGGAGCCCTCTTACGATCAGACGAGGCAACAACCTCCGCTCA
QY      122 Ser-Arg-----ProGluProAsnArgCysGlySerCysArgLysArg
      1991 ACCTCGGCAAGTACCTCTCGAGACCGGAGGTGATGCGCTGCG
      DB
      139 rGlyPheLysCysArgCys-GlyAsnLeuTyrCysAlaLeuHisArgTyr
      2051 ATCTCTGCGG---CACTGCGGAACCTTTACGTCAAGTCTCGCCGCT
      DB
      156 --SerAspLysHisThrCysThrTyrAspTyrLysAlaAlaGly 169
      2108 CTTCCGATACCCCGCGCGCCAGCTTTGAAGCGCGCCGATGCG 2151
      DB

RESULT 14
US-10-821-234-165/c
; Sequence 165, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pc_seq_genes Version 1.0
; SEQ ID NO 165
; LENGTH: 2815
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-821-234-165

```

Alignment Scores:

```

Pred. No.:      489      Length:      2815
Score:          83.00     Matches:      43
Percent Similarity: 37.06%  Conservative: 20
Best Local Similarity: 25.29%  Mismatches: 66
Query Match:      8.43%     Indels:      41
DB:              6         Gaps:      8

```

US-10-716-089-20 (1-188) x US-10-821-234-165 (1-2815)

```

QY 22 LysanLeuCyGlyPheheGlySerGlnAlaThrMeGlyLeuCySerIysCySerTyr 41
   |||||
Db 1802 AAAAACATTGCTGAGCTTCAAGCTCCAGACACACACTCAAGGCTCTGGGGAGGCC-- 1746
   |||||
QY 42 ArgGluThrValMetGlnAlaLeuMetThrAla-----LeuAlaGluGlnAla 57
   |||||
Db 1745 -----CGATGCTCAATGCCAGCATTTGGCATCTTCGTGAGACTACCGGTGCCAGC 1692
   |||||
QY 58 ThrGlnAlaAlaGlnAlaThrSerAlaThrAlaAlaAlaValGlnProProAlaProVal 77
   |||||
Db 1691 ATTGCTCACTTCATGAGCTCCAGCTCGGTGAGGCTGACAGCTCAACCGCAGATC 1632
   |||||
QY 78 HisGluThr-----LysLeuThrCySerGluValGluArgThrMetLeuAla-ProHisGlu 95
   |||||
Db 1631 CAGACCACTCCCTCCCAAGACACTTGTAGCCCGAGGCTGCTCATGCTCTGTATCCAGACA 1572
   |||||
QY 95 nSerSerSerTyrGlnGlnAlaPheLeuValThrProAlaAlaAlaAlaProGlnAlaValIly 115
   |||||
Db 1571 AAGT-----GTGCCAGACAGCCATGACCTGCTGACTATAAG 1536
   |||||
QY 115 sSerSerIle-----AlaAlaProSerArgProGluProAlaValArgCySerIys 131
   |||||
Db 1535 AAAGAGCTTGTGGCGCATTTGAAGTGACATCAAGGCTGACCTTGTGGAGACCCAGC 1476
   |||||
QY 131 rCyArgGluValArgValGlyLeuThrGlyPheIys----- 142
   |||||
Db 1475 TGTGAGGCGAGAGACAGAGGGGTGACAAATTGTCTTGGCGAGTCTTACTCCAGCTGTGGC 1416
   |||||
QY 143 -CyArgCySv-----GlyAlaLeuTyrCyValAlaLeuIly 153
   |||||
Db 1415 ATGCAGGTGTACAGCAAGTATGATCAAGCATGAGGGGTGTCAATATCTGTGAGCTCA 1356
   |||||
QY 153 sArgTyrSerArgIlyHisThrCySerThr 162
   |||||
Db 1355 TCACCAACAGCGGAAAAAGG---TGCACCT 1331

```

RESULT 15

```

US-10-508-263-103
/ Sequence 103, Application US/10508263
/ Publication No. US20050260754A1
/ GENERAL INFORMATION:
/ APPLICANT: BASF Plant Science GmbH
/ TITLE OF INVENTION: Constructs and methods for regulating gene expression
/ FILE REFERENCE: 53262-20085.00
/ CURRENT APPLICATION NUMBER: US/10/508,263
/ NUMBER OF SEQ ID NOS: 126
/ SOFTWARE: Patencin Ver. 2.1
/ SEQ ID NO 103
/ LENGTH: 801
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (1)..(798)
/ OTHER INFORMATION: 22kd alpha-zelin
US-10-508-263-103

```

Alignment Scores:

```

Pred. No.: 115 Length: 801
Score: 82.50 Matches: 48
Percent Similarity: 28.89% Conservative: 7
Best Local Similarity: 26.09% Mismatches: 91
Query Match: 8.38% Indels: 38
DB: 6 Gaps: 6

```

US-10-716-089-20 (1-188) x US-10-508-263-103 (1-801)

```

QY 32 AlaThrMetGlyLeuCySerIysCySerTyrArgGluThrValMetGlnAlaIysMetThr 51
   |||||
Db 168 GCTACAAACAGCGCTTGGCGAGCGCTTTACACAAACCGTTTGGCCCATTTACAAACACA 227
   |||||

```

```

QY 52 AlaLeu-----AlaGluGlnAlaThrGlnAlaAlaGlnAlaAlaThrSerAlaI 70
   |||||
Db 228 ATCTTGGCACAATGTAACATTAACAACATCGCAACAACTAAGGCAAC 287
   |||||
QY 71 ValGlnProProAlaProValHisGluThrIlyValThrCySerGluValG 90
   |||||
Db 288 CGCATTAAGCCCAAGTACGCGGGTGAACCTGTGTCTTAATTGCAAGAC 347
   |||||
QY 91 IleValProHisGlnSerSerSerTyrGlnGlnIlePheLeuValThrPro 110
   |||||
Db 348 ATCCAAACCACTTGTCTGGGGAACACAGCCGC-----ATACAG 398
   |||||
QY 111 ProGlnAlaValIlySerSerIleAlaAlaProSerArgProGlu----- 125
   |||||
Db 399 GTTGACAGATTTCATACAGCTTTATGCAACTAATCCAGGGTAAACCT 458
   |||||
QY 126 -----ProAlaValArgCySerCyS- 132
   |||||
Db 459 GCACAGCAACAATGTTTCATGTAGCCCACTGCTGTGGCAATGCG 518
   |||||
QY 133 -----ArgIlyArgValGlyLeuThrGlyPheIysC 146
   |||||
Db 519 GCACAGCAAGCTGTACAAACAGATCGTACCGGC-----TCTTAGTCAG 572
   |||||
QY 147 AsnLeuTyrCyValAlaLeuHisArgTyrSerAsp----- 159
   |||||
Db 573 GAACCTTACTGCTACTTAACAACAGCTTCTTCATTCACAACATAGAT 632
   |||||
QY 160 ThrCySerThrTyrAspTyrIlyValAlaGluGln-GluAlaIleAlaIly 179
   |||||
Db 633 TGCTGCTAGCTTACAAACAGCGGACAGACTATTAACTCACTTGACAGCG 692
   |||||
QY 179 uValValAla 182
   |||||
Db 693 GGTGCGCCGCC 702

```

Search completed: December 8, 2005, 19:02:26
 Job time : 265 secs